



FORM NO. 1 (Rev. 11-2000) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER <b>PX3813USW</b>
<b>TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371</b>		U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR <b>10/018878</b>
INTERNATIONAL APPLICATION NO <b>PCT/GB00/02357</b>	INTERNATIONAL FILING DATE <b>30 June 2000</b>	PRIORITY DATE CLAIMED <b>1 July 1999</b>
TITLE OF INVENTION <b>VECTORS, CELLS AND PROCESSES FOR PYRIMIDINE DEOXYRIBONUCLEOSIDES PRODUCTION</b>		
APPLICANT(S) FOR DO/EO/US <b>ANDERSON, David Martin; LIU, Lin; PODKOVRV, Sergey; WANG, Baomin</b>		
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:		
<ol style="list-style-type: none"> <li>1. <input checked="" type="checkbox"/> This is a <b>FIRST</b> submission of items concerning a filing under 35 U.S.C. 371.</li> <li>2. <input type="checkbox"/> This is a <b>SECOND</b> or <b>SUBSEQUENT</b> submission of items concerning a filing under 35 U.S.C. 371.</li> <li>3. <input checked="" type="checkbox"/> This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (24) indicated below.</li> <li>4. <input type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (Article 31).</li> <li>5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371 (c) (2))           <ol style="list-style-type: none"> <li>a. <input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau).</li> <li>b. <input checked="" type="checkbox"/> has been communicated by the International Bureau.</li> <li>c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).</li> </ol> </li> <li>6. <input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371 (c) (2)).           <ol style="list-style-type: none"> <li>a. <input type="checkbox"/> is attached hereto.</li> <li>b. <input type="checkbox"/> has been previously submitted under 35 U.S.C. 154(d)(4).</li> </ol> </li> <li>7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c) (3))           <ol style="list-style-type: none"> <li>a. <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau).</li> <li>b. <input type="checkbox"/> have been communicated by the International Bureau.</li> <li>c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</li> <li>d. <input checked="" type="checkbox"/> have not been made and will not be made.</li> </ol> </li> <li>8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c) (3)).</li> <li>9. <input type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).</li> <li>10. <input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).</li> <li>11. <input checked="" type="checkbox"/> A copy of the International Preliminary Examination Report (PCT/IPEA/409).</li> <li>12. <input checked="" type="checkbox"/> A copy of the International Search Report (PCT/ISA/210).</li> </ol>		
<b>Items 13 to 20 below concern document(s) or information included:</b>		
<ol style="list-style-type: none"> <li>13. <input checked="" type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.</li> <li>14. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.</li> <li>15. <input checked="" type="checkbox"/> A <b>FIRST</b> preliminary amendment.</li> <li>16. <input type="checkbox"/> A <b>SECOND</b> or <b>SUBSEQUENT</b> preliminary amendment.</li> <li>17. <input type="checkbox"/> A substitute specification.</li> <li>18. <input type="checkbox"/> A change of power of attorney and/or address letter.</li> <li>19. <input checked="" type="checkbox"/> A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825.</li> <li>20. <input type="checkbox"/> A second copy of the published international application under 35 U.S.C. 154(d)(4).</li> <li>21. <input type="checkbox"/> A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).</li> <li>22. <input checked="" type="checkbox"/> Certificate of Mailing by Express Mail</li> <li>23. <input checked="" type="checkbox"/> Other items or information:</li> </ol>		
<b>Sequence Listing (paper copy) and Statement of Support PCT publication cover page PCT Request</b>		

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR 1.53) <b>10/018878</b>		INTERNATIONAL APPLICATION NO. <b>PCT/GB00/02357</b>		ATTORNEY'S DOCKET NUMBER <b>PX3813USW</b>	
24. The following fees are submitted: <b>BASIC NATIONAL FEE ( 37 CFR 1.492 (a) (1) - (5) ) :</b>				<b>CALCULATIONS PTO USE ONLY</b>	
<input type="checkbox"/> Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO .....				<b>\$1040.00</b>	
<input checked="" type="checkbox"/> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO .....				<b>\$890.00</b>	
<input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO .....				<b>\$740.00</b>	
<input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) .....				<b>\$710.00</b>	
<input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) .....				<b>\$100.00</b>	
<b>ENTER APPROPRIATE BASIC FEE AMOUNT =</b>				<b>\$890.00</b>	
Surcharge of <b>\$130.00</b> for furnishing the oath or declaration later than months from the earliest claimed priority date (37 CFR 1.492 (e)). <input type="checkbox"/> 20 <input type="checkbox"/> 30				<b>\$0.00</b>	
<b>CLAIMS</b>	<b>NUMBER FILED</b>	<b>NUMBER EXTRA</b>	<b>RATE</b>		
Total claims	48    - 20 =	28	x    \$18.00	<b>\$504.00</b>	
Independent claims	3    - 3 =	0	x    \$84.00	<b>\$0.00</b>	
Multiple Dependent Claims (check if applicable).				<input type="checkbox"/>	<b>\$0.00</b>
<b>TOTAL OF ABOVE CALCULATIONS =</b>				<b>\$1,394.00</b>	
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27). The fees indicated above are reduced by 1/2.				<b>\$0.00</b>	
<b>SUBTOTAL =</b>				<b>\$1,394.00</b>	
Processing fee of <b>\$130.00</b> for furnishing the English translation later than months from the earliest claimed priority date (37 CFR 1.492 (f)). <input type="checkbox"/> 20 <input type="checkbox"/> 30				<b>\$0.00</b>	
<b>TOTAL NATIONAL FEE =</b>				<b>\$1,394.00</b>	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) (check if applicable). <input type="checkbox"/>				<b>\$0.00</b>	
<b>TOTAL FEES ENCLOSED =</b>				<b>\$1,394.00</b>	
				<b>Amount to be: refunded</b>	<b>\$</b>
				<b>charged</b>	<b>\$</b>
a. <input type="checkbox"/> A check in the amount of _____ to cover the above fees is enclosed. b. <input checked="" type="checkbox"/> Please charge my Deposit Account No. <u>07-1392</u> in the amount of <b>\$1,394.00</b> to cover the above fees. A duplicate copy of this sheet is enclosed. c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. <u>07-1392</u> A duplicate copy of this sheet is enclosed. d. <input type="checkbox"/> Fees are to be charged to a credit card. <b>WARNING:</b> Information on this form may become public. <b>Credit card information should not be included on this form.</b> Provide credit card information and authorization on PTO-2038.					
<b>NOTE:</b> Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.					
SEND ALL CORRESPONDENCE TO:					
<div style="border: 1px solid black; padding: 10px; width: 150px; margin: 0 auto;">   <b>23347</b>  <small>PATENT TRADEMARK OFFICE</small> </div>					
<div style="border-bottom: 1px solid black; width: 200px; margin: 0 auto;">   <small>SIGNATURE</small> </div> <div style="margin-top: 10px;"> <b>Frank P. Grassler</b>  <small>NAME</small>  <b>31,164</b>  <small>REGISTRATION NUMBER</small>  <b>December 19, 2001</b>  <small>DATE</small> </div>					

<b>CERTIFICATE OF MAILING BY "EXPRESS MAIL" (37 CFR 1.10)</b>		Docket No.
Applicant(s): ANDERSON et al.		PX3813USW

Serial No. To be assigned	Filing Date Concurrently herewith	Examiner	Group Art Unit
------------------------------	--------------------------------------	----------	----------------

Invention: **VECTORS, CELLS AND PROCESSES FOR PYRIMIDINE DEOXYRIBONUCLEOSIDES PRODUCTION**

I hereby certify that the following correspondence:

Transmittal Letter concerning a filing under 35 USC 371 and accompanying paperwork

*(Identify type of correspondence)*

is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 in an envelope addressed to: The Assistant Commissioner for Patents, Washington, D.C. 20231 on

12/19/2001  
*(Date)*

Ban Younan

*(Typed or Printed Name of Person Mailing Correspondence)*

Ban Younan

*(Signature of Person Mailing Correspondence)*

EL395895129US

*(Express Mail Tracking Number)*

**Note: Each paper must have its own certificate of mailing.**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

#4/a

In re application of: ANDERSON et al.

International Application No. PCT/GB00/02357

International Filing Date: 30 June 2000

For: *VECTORS, CELLS AND PROCESSES FOR PYRIMIDINE  
DEOXYRIBONUCLEOSIDES PRODUCTION*

BOX PCT (DO/EO/US)  
Commissioner For Patents  
Washington, D.C. 20231

FIRST PRELIMINARY AMENDMENT

Sir:

The above-identified application is being transmitted herewith for entry into the U.S. National Phase under Chapter II of the PCT. For the purposes of adding the priority information, please amend the application as follows:

**In the Abstract:**

Please substitute the attached Abstract, which has been placed on a separate piece of paper according to US practice.

**In the Specification:**

On the first line of the specification, after the Title, please add:

--This application is filed pursuant to 35 U.S.C. § 371 as a United States National Phase Application of International Application No. PCT/GB00/02357 filed June 30, 2000, which claims priority from U.S. Serial No. 09/345,492 filed July 1, 1999, and U.S. Serial No. 60/141,827 filed July 1, 1999.--

Clean Copy of Pending Claims

Claims

1. A DNA construct comprising a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene or a uridine kinase gene and/or a dCTP deaminase gene.
2. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene.
3. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a uridine kinase gene and/or a dCTP deaminase gene.
4. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a uridine kinase gene and a dCTP deaminase gene.
5. A DNA construct according to claim 1 wherein the construct is an extrachromosomal vector.
6. A DNA construct according to claim 5 wherein the vector is a plasmid, virus, transposon, minichromosome or phage.
7. A DNA construct according to claim 6 wherein the vector is a plasmid, wherein the ribonucleotide reductase gene is a *nrdA* gene and/or a *nrdB* gene, and wherein the thioredoxin gene is a *nrdC* gene and said genes are arranged in an operon.
8. A DNA construct according to claim 7 wherein the transcriptional unit comprises a *nrdA* gene and a *nrdB* gene.
9. A DNA construct according to claim 8 in which the *nrdA* and *nrdB* genes are located downstream of the *nrdC* gene.

10. A DNA construct according to claim 8 wherein the *nrdC* gene is upstream of the *nrdA* gene and *nrdA* gene is upstream of the *nrdB* gene.
11. A DNA construct according to claim 6, further comprising a regulatory element.
12. A DNA construct according to claim 11, wherein the regulatory element is selected from the group consisting of a promoter, an operator, a termination sequence, an initiation sequence and a ribosome binding site.
13. A DNA construct according to claim 12 wherein the promoter is the lambda P<sub>L</sub> promoter or a derivative therefrom.
14. A DNA construct according to claim 12 wherein the termination sequence is a synthetic terminator.
15. A DNA construct according to claim 7 wherein the *nrdA* gene is modified such that ribonucleotide reductase encoded by the unit is less sensitive to allosteric inhibition than the ribonucleotide reductase encoded by the unit comprising an unmodified *nrdA* gene.
16. A DNA construct according to claim 15, wherein the *nrdA* gene is modified at a dTTP binding site.
17. A DNA construct according to claim 15 wherein the *nrdA* gene is modified so as to encode an Ala-to-Ile change at position 79 in the *nrdA* expression product.
18. A DNA construct according to claim 7 wherein the *nrdA*, *nrdB* and *nrdC* genes are derived from a T phage.

19. A DNA construct according to claim 18 wherein the T phage is a T even phage.
20. A DNA construct according to claim 19 wherein the T even phage is a T4 phage.
21. A DNA construct according to claim 1 wherein the construct further comprises a thymidylate synthase gene.
22. A DNA construct according to claim 21 wherein the thymidylate synthase gene is the *td* gene.
23. A DNA construct according to claim 22 wherein *td* gene is located on the same vector as the *nrdA*, *nrdB* and *nrdC* genes.
24. A DNA construct according to claim 22 wherein the *td* gene is located in the same operon as the *nrdA*, *nrdB* and *nrdC* genes.
25. A DNA construct according to claim 24 wherein the *td* gene is located downstream (in terms of reading frame) from the *nrdA*, *nrdB* and *nrdC* genes.
26. A DNA construct according to claim 22 wherein the *td* gene is derived from a T phage.
27. A DNA construct according to claim 26 wherein the *td* gene is derived from a T even phage.
28. A DNA construct according to claim 27 wherein the *td* gene is derived from T4 phage.

29. A DNA construct according to claim 2 further comprising a uridine kinase gene and/or a dCTP deaminase gene.

30. A DNA construct according to claim 29, wherein the DNA construct comprises both a uridine kinase gene and a dCTP deaminase gene.

31. A DNA construct according to claim 1, wherein the uridine kinase gene is a *udk* gene.

32. A DNA construct according to claim 1, wherein the dCTP deaminase gene is a *dcd* gene.

33. A modified host cell comprising a DNA construct according to any of the preceding claims.

34. A modified host cell according to claim 33 wherein the host cell displays repressed or no uracil DNA glycosylase activity.

35. A modified host cell according to claim 34 wherein one or more host cell DNA sequences encoding for uracil DNA glycosylase activity have been modified so as to encode expression products displaying repressed, low levels of or no uracil DNA glycosylase activity.

36. A modified host cell according to claim 35 wherein the host cell DNA sequence is an *ung* gene.

37. A modified host cell according to claim 34 wherein one or more host cell DNA sequences encoding for uracil DNA glycosylase activity have been removed.



38. A modified host cell according to claim 33 wherein the cell is a eukaryote or a prokaryote.

39. A modified host cell according to claim 38 wherein the host cell is selected from the group consisting of *E. coli*, *Salmonella*, *Pseudomonas*, *Bacillus* and *Saccharomyces*.

40. A modified host cell comprising a DNA construct, which construct comprises a transcription DNA unit, which unit comprises a ribonucleotide reductase gene and a thioredoxin gene, wherein said ribonucleotide reductase displays less sensitivity to allosteric inhibition than the wild-type equivalent of the reductase, and wherein said host cell further comprises one or more of the following features:

- (a) a transcription unit located on said DNA construct, comprising a thymidylate synthase gene heterologous to the thymidylate synthase gene of the host cell;
- (b) a transcription unit located on said DNA construct, comprising a uridine kinase gene;
- (c) a transcription unit located on said DNA construct, comprising a dCTP deaminase gene; and
- (d) repressed or absent uracil DNA glycosylase activity.

41. A modified host cell according to claim 40, wherein the ribonucleotide reductase gene is modified at a dTTP binding site.

42. A modified host cell according to claim 41, wherein the ribonucleotide reductase is the *nrdA* gene and the modification is an Ala-to-Ile change at position 79 in the *nrdA* expression product.

43. A modified host cell according to claim 38, wherein the DNA construct comprises both the uridine kinase gene and the dCTP deaminase gene.

44. A modified host cell according to claim 40, wherein the cell comprises each one of the features of (a) to (d).

45. A process for the production of pyrimidine deoxyribonucleosides comprising culturing a host cell according to claim 33.

46. A process according to claim 45 wherein the deoxyribonucleoside is thymidine.

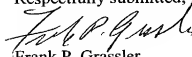
47. A process for the manufacture of azidothymidine comprising culturing a host cell according to claim 33.

48. A culture medium comprising a host cell according to claim 33.

**REMARKS**

Currently claims 1-48 are pending. Claims 5, 21, 31-32, 45, and 47-48 have been amended to place them in form appropriate to US practice and to reduce the filing fee by removing multiple dependency. Attached hereto is a marked-up version of the changes made to the claims by the current amendment. The attached page is captioned "**Version With Markings To Show Changes Made**". Applicants have attached an abstract on a separate sheet of paper as required by US practice. Applicants have amended the specification for purposes of adding the priority information.

Respectfully submitted,

  
Frank P. Grassler  
Attorney for Applicants  
Registration No. 31,164

Date: 12-18-01  
GlaxoSmithKline Inc.  
Corporate Intellectual Property  
Five Moore Drive, P.O. Box 13398  
Research Triangle Park, NC 27709  
Phone: 919-483-2482  
Facsimile: 919-483-7988

**Abstract****VECTORS, CELLS AND PROCESSES FOR PYRIMIDINE  
DEOXYRIBONUCLEOSIDES PRODUCTION**

Novel DNA constructs and host cells comprising the same are disclosed. DNA constructs comprise a transcription unit ( e.g. operon) comprising DNA sequences encoding for ribonucleotide reductase and thioredoxin or a uridine kinase gene and/or a dCTP deaminase gene. In preferred embodiments the constructs comprising DNA sequences encoding for ribonucleotide reductase and thioredoxin further comprise DNA sequences encoding for thymidylate synthase and/or transcription units comprising sequences encoding for uridine kinase preferably together with dCTP deaminase. In particularly preferred embodiments, the host cells comprise constructs having all of the above characteristics wherein the host cell displays repressed or no uracil DNA glycosylase activity. This may be achieved by removal of the host cell *ung* gene. Use of host cells in the manufacture of pyrimidine deoxyribonucleotides e.g. thymidine is also disclosed.

Version With Markings To Show Changes Made

Claims

1. A DNA construct comprising a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene or a uridine kinase gene and/or a dCTP deaminase gene.
2. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene.
3. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a uridine kinase gene and/or a dCTP deaminase gene.
4. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a uridine kinase gene and a dCTP deaminase gene.
5. A DNA construct according to claim 1 [any of claims 1-4] wherein the construct is an extrachromosomal vector.
6. A DNA construct according to claim 5 wherein the vector is a plasmid, virus, transposon, minichromosome or phage.
7. A DNA construct according to claim 6 wherein the vector is a plasmid, wherein the ribonucleotide reductase gene is a *nrdA* gene and/or a *nrdB* gene, and wherein the thioredoxin gene is a *nrdC* gene and said genes are arranged in an operon.
8. A DNA construct according to claim 7 wherein the transcriptional unit comprises a *nrdA* gene and a *nrdB* gene.
9. A DNA construct according to claim 8 in which the *nrdA* and *nrdB* genes are located downstream of the *nrdC* gene.

10. A DNA construct according to claim 8 wherein the *nrdC* gene is upstream of the *nrdA* gene and *nrdA* gene is upstream of the *nrdB* gene.
11. A DNA construct according to claim 6, further comprising a regulatory element.
12. A DNA construct according to claim 11, wherein the regulatory element is selected from the group consisting of a promoter, an operator, a termination sequence, an initiation sequence and a ribosome binding site.
13. A DNA construct according to claim 12 wherein the promoter is the lambda P<sub>L</sub> promoter or a derivative thereof.
14. A DNA construct according to claim 12 wherein the termination sequence is a synthetic terminator.
15. A DNA construct according to claim 7 wherein the *nrdA* gene is modified such that ribonucleotide reductase encoded by the unit is less sensitive to allosteric inhibition than the ribonucleotide reductase encoded by the unit comprising an unmodified *nrdA* gene.
16. A DNA construct according to claim 15, wherein the *nrdA* gene is modified at a dTTP binding site.
17. A DNA construct according to claim 15 wherein the *nrdA* gene is modified so as to encode an Ala-to-Ile change at position 79 in the *nrdA* expression product.
18. A DNA construct according to claim 7 wherein the *nrdA*, *nrdB* and *nrdC* genes are derived from a T' phage.

19. A DNA construct according to claim 18 wherein the T phage is a T even phage.
20. A DNA construct according to claim 19 wherein the T even phage is a T4 phage.
21. A DNA construct according to claim 1 [any of claims 1-4] wherein the construct further comprises a thymidylate synthase gene.
22. A DNA construct according to claim 21 wherein the thymidylate synthase gene is the *td* gene.
23. A DNA construct according to claim 22 wherein *td* gene is located on the same vector as the *nrdA*, *nrdB* and *nrdC* genes.
24. A DNA construct according to claim 22 wherein the *td* gene is located in the same operon as the *nrdA*, *nrdB* and *nrdC* genes.
25. A DNA construct according to claim 24 wherein the *td* gene is located downstream (in terms of reading frame) from the *nrdA*, *nrdB* and *nrdC* genes.
26. A DNA construct according to claim 22 wherein the *td* gene is derived from a T phage.
27. A DNA construct according to claim 26 wherein the *td* gene is derived from a T even phage.
28. A DNA construct according to claim 27 wherein the *td* gene is derived from T4 phage.

29. A DNA construct according to claim 2 further comprising a uridine kinase gene and/or a dCTP deaminase gene.

30. A DNA construct according to claim 29, wherein the DNA construct comprises both a uridine kinase gene and a dCTP deaminase gene.

31. A DNA construct according to claim 1,[3 or 4] wherein the uridine kinase gene is a *udk* gene.

32. A DNA construct according to claim 1,[3 or 4] wherein the dCTP deaminase gene is a *dcd* gene

33. A modified host cell comprising a DNA construct according to any of the preceding claims.

34. A modified host cell according to claim 33 wherein the host cell displays repressed or no uracil DNA glycosylase activity.

35. A modified host cell according to claim 34 wherein one or more host cell DNA sequences encoding for uracil DNA glycosylase activity have been modified so as to encode expression products displaying repressed, low levels of or no uracil DNA glycosylase activity.

36. A modified host cell according to claim 35 wherein the host cell DNA sequence is an *ung* gene.

37. A modified host cell according to claim 34 wherein one or more host cell DNA sequences encoding for uracil DNA glycosylase activity have been removed.

38. A modified host cell according to claim 33 wherein the cell is a eukaryote or a prokaryote.

39. A modified host cell according to claim 38 wherein the host cell is selected from the group consisting of *E. coli*, *Salmonella*, *Pseudomonas*, *Bacillus* and *Saccharomyces*.

40. A modified host cell comprising a DNA construct, which construct comprises a transcription DNA unit, which unit comprises a ribonucleotide reductase gene and a thioredoxin gene, wherein said ribonucleotide reductase displays less sensitivity to allosteric inhibition than the wild-type equivalent of the reductase, and wherein said host cell further comprises one or more of the following features:

- (a) a transcription unit located on said DNA construct, comprising a thymidylate synthase gene heterologous to the thymidylate synthase gene of the host cell;
- (b) a transcription unit located on said DNA construct, comprising a uridine kinase gene;
- (c) a transcription unit located on said DNA construct, comprising a dCTP deaminase gene; and
- (d) repressed or absent uracil DNA glycosylase activity.

41. A modified host cell according to claim 40, wherein the ribonucleotide reductase gene is modified at a dTTP binding site.

42. A modified host cell according to claim 41, wherein the ribonucleotide reductase is the *nrda* gene and the modification is an Ala-to-Ile change at position 79 in the *nrda* expression product.

43. A modified host cell according to claim 38, wherein the DNA construct comprises both the uridine kinase gene and the dCTP deaminase gene.



44. A modified host cell according to claim 40, wherein the cell comprises each one of the features of (a) to (d).

45. A process for the production of pyrimidine deoxyribonucleosides comprising culturing a host cell according to claim 33 [any of claims 33-44].

46. A process according to claim 45 wherein the deoxyribonucleoside is thymidine.

47. A process for the manufacture of azidothymidine comprising culturing a host cell according to claim 33 [any of claims 33-44].

48. A culture medium comprising a host cell according to claim 33 [any of claims 33-44].

**Vectors, Cells and Processes for  
Pyrimidine Deoxyribonucleosides Production**

**Field of the Invention**

5 The present invention relates to the production of pyrimidines, purines and derivatives thereof e.g. deoxyribonucleosides, using genetically modified cells comprising novel DNA constructs.

**Background of the invention**

10 Thymidine is useful as a pharmaceutical intermediate, particularly for the chemical synthesis of azidothymidine ("AZT," sold under the trademark ZIDOVUDINE). Although ZIDOVUDINE-type AZT was one of the first therapies developed for HIV/AIDS, it continues to have important and expanded use (Langreth, R., *The Wall*  
15 *Street Journal*, Nov. 21, 1995, pp B12). ZIDOVUDINE-type AZT is valuable particularly when used in combination therapies such as a combination with lamivudine (also known as 3TC), sold under the trademark EPIVIR. This lamivudine and 3TC  
20 combination is sold under the trademark COMBIVIR. Although the HIV virus can mutate to form resistance to either AZT or 3TC, COMBIVIR-type nucleotide-analog combination is particularly effective because the reverse transcriptase apparently  
25 cannot be resistant to both nucleoside analogues at the same time (Larder, B.A. *et al.*, *Science* 269: 696-699, 1995). ZIDOVUDINE-type AZT is also useful in conjunction with HIV protease inhibitor type drugs (Waldholz, M., *The Wall Street Journal*, Jan. 30, 1996, pp B1), and in the treatment of HIV infected pregnant women in order to reduce  
30 the frequency of infection of the fetus at birth. In 1997 about 600,000 children died of AIDS contracted from their mothers at birth. ZIDOVUDINE-type AZT taken for several months prior to birth can reduce the transmission of the virus to infants by two-thirds. Thymidine produced by chemical synthesis used in the manufacture of AZT is a very significant cost.

In U.S. 5,213,972 (McCandliss & Anderson, hereinafter "the '972 patent"), the entire contents of which are incorporated herein by reference and to which the reader is specifically referred, a process for the production of pyrimidine deoxyribonucleoside

WO 01/02580

PCT/GB00/02357

(PdN) is disclosed (see in particular examples 7 to 14 of the '972 patent). A replicatable microorganism comprising and expressing a DNA sequence encoding a pyrimidine deoxyribonucleotide phosphohydrolase that converts a PdN monophosphate to a pyrimidine deoxyribonucleoside is taught. More particularly, McCandliss & Anderson, *supra*, describe a fermentation method that can be used to produce thymidine that involves the expression of deoxythymidylate phosphohydrolase (dTMPase) from the *Bacillus* bacteriophage PBS1. This type of enzyme has been found in nature expressed by bacteriophages that do not contain thymidine in their DNA, but instead incorporates compounds like deoxyuridine or hydroxymethyl-deoxyuridine.

In the thymidine fermentation described in the '972 patent, the enzymes that degrade thymidine (thymidine phosphorylase and uridine phosphorylase) have been removed by mutation so that thymidine accumulates. Thus, the use of the dTMPase enzyme helps create the pathway to allow thymidine synthesis. An expression of dTMPase alone, however, may not assure a commercially viable level of thymidine production. Accordingly, there is a continuing need to enhance the production of thymidine by cells expressing dTMPase in order to make thymidine production by fermentation commercially viable, by lowering the production cost relative to the current chemical synthesis methods.

The biochemical pathway for pyrimidine deoxynucleotide production, for example, in *E. coli* is highly regulated at the levels of transcription and translation as well as at the protein level by mechanisms including attenuation, feedback inhibition and enzyme activation. Neuhaard, J. and R.A. Kelln, Biosynthesis and Conversion of Pyrimidines, Chapter 35 [In] Neidhardt, F.C. *et al.* [eds] "*Escherichia coli* and *Salmonella* Cellular and Molecular Biology", Second Edition, Vol. I, pp580-599, ASM Press, Washington D.C., 1996. The expression of dTMPase and elimination of thymidine breakdown by mutations in the *deoA* (thymidine phosphorylase), *udp* (uridine phosphorylase) and *tdk* (thymidine kinase) genes and therefore resulting expression products results in thymidine synthesis in *E. coli* but not at a commercially viable level.

**Summary of the invention**

The biosynthesis of purines and pyrimidines involves a common step of reducing a ribonucleoside diphosphate (in some species triphosphate) to its corresponding deoxy analog. In the overall process the reduction of the ribose moiety to 2-deoxyribose requires a pair of hydrogen atoms which are ultimately donated by NADPH and H<sup>+</sup>. However, the immediate electron donor is not NADPH but the reduced form of a heat stable protein called thioredoxin or glutaredoxin and at least one other unidentified source since the *E. coli* ribonucleotide reductase system still works in *trxA* (thioredoxin) *grx* (glutaredoxin) double mutants (Neuhard and Kelln, *supra*). The reducing equivalents of the reduced thioredoxin are transferred to ribonucleoside diphosphate reductase which carries out the reduction process. Manipulation of, for example, this step could prove useful in improving the commercial production of purine and pyrimidine deoxynucleosides.

It is an object of the present invention to provide novel DNA constructs e.g. vectors and genetically modified microorganisms comprising said vectors particularly for use in the production of recoverable amounts, especially commercially useful amounts, of pyrimidine and purine deoxynucleosides.

It is also an object of the present invention to provide processes which represent an improvement over McCandliss and Anderson described *supra*.

In accordance with one aspect of the present invention there is provided a DNA construct comprising a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene or a uridine kinase gene and/or a dCTP deaminase gene.

In one embodiment the DNA construct comprises a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene.

In another embodiment the DNA construct comprises a transcriptional unit which comprises a uridine kinase gene and/or a dCTP deaminase gene.

WO 01/02580

PCT/GB00/02357

Preferably the DNA construct comprises a transcriptional unit which comprises a uridine kinase gene and a dCTP deaminase gene.

Most preferably the DNA construct comprises a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene and a uridine kinase gene and a dCTP deaminase gene.

In accordance with another aspect of the present invention there is provided a modified host cell comprising a DNA construct according to the invention.

In accordance with yet another aspect of the present invention there is provided a culture medium comprising the modified host cells of the invention and processes for the production of a purine or pyrimidine, for example thymidine, comprising the use of said modified host cells.

In one embodiment the host cells comprise a DNA construct which construct comprises a transcription DNA unit (e.g. operon) which unit comprises DNA sequences encoding for ribonucleotide reductase and thioredoxin in which said reductase preferably displays less sensitivity to allosteric inhibition than a wild type host cell equivalent or counterpart wherein said cell further comprises one or more of the following features:

- (a) a transcription unit (e.g. operon), preferably located on said DNA construct, comprising DNA sequences encoding for (and preferably heterologous with respect to host cell equivalent) thymidylate synthase;
- (b) a transcription unit (e.g. operon), preferably located on said DNA construct, comprising DNA sequences encoding for uridine kinase and preferably dCTP deaminase; and
- (c) repressed or absent Uracil DNA glycosylase activity.

In another embodiment the DNA construct for use in the production of recoverable amounts of pyrimidine and derivatives thereof, in particular pyrimidine deoxyribonucleosides such as thymidine, comprises a transcription unit (e.g. operon) which unit comprises (preferably heterologous) DNA sequences encoding for uridine kinase and/or dCTP deaminase.

Genetically modified host cells comprising and expressing the construct and culture medium comprising the modified host cells are also provided.

5 This aspect is based, in part, on the observation that host cells comprising DNA encoding for uridine kinase and/or dCTP deaminase, optionally together with additional genes as suggested in US 5,213,972 required for thymidine production, lead to a significant improvement in thymidine production.

10 The respective aspects of the present invention disclose for the first time a plurality of advances on the teaching of U.S 5,213,972 to provide improved DNA constructs and host cells comprising the constructs for use in the commercial production of pyrimidine deoxyribonucleosides, particularly thymidine.

15 Other objects, features and advantages of the present invention will become apparent from the following description. It should be understood, however, that these represent preferred embodiments of the invention and are by way of illustration only. Various modifications and changes within the spirit and scope of the invention will become apparent to those skilled in the art.

### **Preferred embodiments of the invention**

The construct of the present invention may be chromosomal or more preferably extra-chromosomal e.g. located on a vector.

25 Vectors of the present invention include plasmid, virus, transposons, minichromosome or phage, preferably plasmid. The vector comprising the transcription unit may be introduced into the host cell according to any convenient method known to those skilled in the art, e.g. P1 transduction, electroporation or transformation. Suitable host cells  
30 useful in the present invention include eukaryotes and prokaryotes (e.g. Bacterium). Prokaryotes include *E. coli*, *Salmonella*, *Pseudomonas*, *Bacillus*, strains and mutants thereof. *E. coli* is preferred due to the large amount of information, genetic tools and mutant alleles that are available. It is particularly preferred that a method of

transduction is available for the host cell of choice to enable mutations to be readily moved from one host cell to another and facilitate genetic mutation of the host without requiring direct mutation whenever a new mutation is desired.

- 5 The present inventors have found that the use of bacteriophage T4 *nrdA*, *nrdB* and *nrdC* genes are particularly useful for encoding the reductase and thioredoxin in *E. coli*. See Sjöberg, B. M. *et al.*, *EMBO J.*, 5:2031-2036 (1986); Tseng, M.-J., *et al.*, *J. Biol. Chem.* 263:16242-16251 (1988); and LeMaster, D.M., *J. Virol.* 59:759-760 (1986). More specifically, a very significant improvement in *E. coli* thymidine production was
- 10 achieved through the cloning and expression of the T4 bacteriophage *nrdA* and *nrdB* genes coding for ribonucleotide reductase together with T4 *nrdC* coding for thioredoxin since the T4 ribonucleotide reductase cannot use *E. coli* thioredoxin. The T4-coded ribonucleotide reductase was found to be relatively insensitive to control by allosteric inhibition *in vitro* compared to the *E. coli* enzyme (Berglund, O., *J. Biol. Chem.*
- 15 247:7276-7281, 1972). For example, unlike the *E. coli* enzyme (Berglund, O., *J. Biol. Chem.* 247: 270-7275, 1972) the T4 ribonucleotide reductase is not inhibited by dATP, but actually stimulated by dATP and ATP (Berglund, O., *J. Biol. Chem.* 247:7276-7281, 1972).
- 20 DNA sequences encoding for the ribonucleotide reductase (e.g. *nrdA* and *nrdB* genes) and thioredoxin (e.g. *nrdC* gene) are preferably heterologous with respect to host cell DNA and preferably derived from T phage (preferably *E. coli* T bacteriophage), particularly T "even" phages e.g. T2, T4 or T6. See Campbell, A.M., Bacteriophages, Chapter 123, *In* Neidherdt, *supra*; and Mathews, C.K. *et al.* (eds.) Bacteriaophage T4,
- 25 American Society of Microbiology, Washington, D.C., 1983. The term "derived from" is intended to define not only a source in the sense of its physical origin but also to define material which has structural and/or functional characteristics which correspond to material originating from the reference source.
- 30 Another useful feature of the T even phage enzyme is its substrate specificity. The normal *E. coli* ribonucleotide reductase uses UDP as a substrate only poorly since the  $K_m$  for UDP is about 10 fold higher for UDP than CDP (Neuhard and Kelln, *supra*). However, the T4 enzyme has only a two-fold difference in  $K_m$  (Berglund, O., *J. Biol.*

WO 01/02580

PCT/GB00/02357

Chem. 247: 7276-7281, 1972) between CDP and UDP substrates allowing two routes to dUTP synthesis. Although there have been attempts to obtain functional expression of T4 ribonucleotide reductase in *E. coli*, previous efforts were only successful in expressing the components separately and could demonstrate activity only by mixing *in vitro* (Tseng, M.-J., P. He, J.M. Hilfinger, and G.R. Greenberg, *J. Bacteriol.* 172: 6323-6332, 1990). Whilst not being bound by theory, the inventors believe that perhaps due to the lack of the usual pattern of feedback inhibition, expression of T4 ribonucleotide reductase in *E. coli* is lethal and it must be carefully conditionally expressed. Further envisaged are genes that encode precursor forms of the reductase and/or thioredoxin which are processed to produce a mature form. Such processing may proceed via various intermediate forms.

Vectors of the present invention preferably comprise a regulatory element ( e.g. promoter such as lambda P<sub>L</sub>, operator, activator, repressor such as lambda repressor, particularly a temperature sensitive variant, and/or enhancer), appropriate termination sequences initiation sequences and ribosome binding sites. The vector may further comprise a selectable marker. Alternatively, regulatory elements (particularly lambda repressor) may be located on the host cell chromosome. It is preferred that *nrdA* and *nrdB* are arranged in the vector downstream (in terms of reading frame) from *nrdC*. In particular, it is preferred that *nrdB* is arranged downstream from *nrdA*. Thus a most preferred arrangement is a vector comprising an operon comprising *nrdCAB*.

The T4 ribonucleotide reductase is not devoid of feedback-control *in vivo* (J. Ji, R.G. Sargent, and C.K. Mathews, *J.Biol.Chem.* 266: 16289-16292, 1991; and Berglund *supra*). To promote ribonucleoside diphosphate reduction further e.g. for thymidine production, the gene coding for the regulatory subunit, *nrdA*, may be modified by, for example, a mutational approach to create an enzyme capable of increased thymidine production due to e.g. a reduced sensitivity to allosteric inhibition for example inhibition by the enzyme's immediate product or inhibition by a product resulting from a downstream event.

In order to construct T4 *nrdA* mutants, site-directed mutagenesis may be used to modify or change (e.g. substitute) gene bases encoding amino acids suspected to alter



WO 01/02580

PCT/GB00/02357

e.g. dTTP binding site involved in allosteric regulation. Analysis of the amino acid sequence of T4 ribonucleotide reductase revealed a segment that appears to fit well with a postulated consensus sequence thought to be involved in dTTP binding (E.M. McIntosh and R.H. Haynes, *Mol. Cell. Biol.* 6: 1711-1721, 1986). Several changes may be made in this region of the T4 ribonucleotide reductase using oligonucleotide-directed mutagenesis. The general approach may be modelled after the effort of More *et al.* (More, J.T., J. M. Ciesla, L.-M. Changchien, G.F. Maley and F. Maley, *Biochemistry* 33: 2104-2112, 1994) to reduce the dTTP binding of deoxycytidylate deaminase. One mutation, <sup>79</sup>Ala to Ile, in the T4 *nrdA* appeared to be very useful. For example, the thymidine productivity of strains containing the <sup>79</sup>Ala to Ile mutant in T4 *nrdA* evaluated by a shake flask fermentation method was significantly increased. As demonstrated below, the present inventors achieved at least 25% increase over the parent strain without this single change.

Although the <sup>79</sup>Ala to Ile is one successful example, those skilled in the art will now realize that many other amino acid changes to this region are now possible to obtain the desired effect, that being to putatively disrupt dTTP binding, but not disrupt the enzyme's basic functionality. For example, substitution of <sup>79</sup>Ala with other amino acids displaying similar side chains to Ile (e.g. leucine, valine) may be utilized. Modifications of position 79 in conjunction with other modifications (e.g. mutations) within the postulated consensus region are also envisaged. Deletion of one or more amino acid positions in the consensus region and introduction of synthetic DNA into the region are other approaches available to those skilled in the art.

In another aspect of the present invention there is provided a host cell comprising a construct which construct (e.g. vector) comprises a transcriptional unit comprising DNA sequences encoding for heterologous ribonucleotide reductase and thioredoxin which reductase is less sensitive to allosteric inhibition than the wild type host cell equivalent or counterpart. It will be apparent to those skilled in the art that determining the relative sensitivity of a candidate heterologous reductase to allosteric inhibition compared to the wild type host cell equivalent is a matter of routine experimentation and observation.

WO 01/02580

PCT/GB00/02357

Transcription units comprising the DNA sequences e.g. *nrdA*, *nrdB* and *nrdC* genes are preferably operons wherein the *nrd* genes are arranged in tandem. This permits transcription of these genes as a single mRNA transcript. In order to minimize unproductive energy expenditure by the host cell and further to minimize plasmid size, it is preferred that the operon contains only genetic sequences required in the encoding of reductase and thioredoxin (including any regulatory or control elements). This may necessitate the removal of superfluous DNA (for example, the unusual intron in the phage T4 *nrdB* gene, Sjöberg, B-M., *et al* EMBO J.5: 2031-2036, 1986).

In other preferred embodiments, vectors of the present invention for use in for example the production of thymidine further comprise DNA sequences encoding for thymidylate synthase (e.g. the *td* gene). See e.g. Chu, F. K. *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3049-3053 (1984); Chu, F. K. *et al.*, *J. Bacteriol.* 169:4368-4375 (1987). The purpose of using this enzyme is to improve control over the levels of deoxyuridine produced and in particular the relative impurity level of deoxyuridine relative to thymidine. The dTMPase enzyme is not completely specific for dTMP. With a higher  $K_m$  than for dTMP, the PBS1 dTMPase will also utilize dUMP as substrate to produce deoxyuridine (Price, A. R., *Methods in Enzymol.* 51: 285-290, 1978). Deoxyuridine creates a significant problem for thymidine purification. Therefore, one way to reduce deoxyuridine production is to efficiently convert dUMP to dTMP by increasing the level or effectiveness of thymidylate synthase such that the internal concentration of dUMP always remains very low.

The thymidylate synthase gene (*td*) may be heterologous with respect to the host cell and it is preferred that *td* is derived from (in the sense defined supra) T bacteriophage, e.g. T "even" phage and in particular T4 phage *td*. Although *td* may be located in its own transcription unit, it is preferred that *td* is located in the same transcription unit e.g. operon as *nrd* genes. Moreover, it is preferred that *td* is located in the same operon downstream (in terms of reading frame) from the *nrd* genes.

McCandliss and Anderson, *supra*, amplified the *E. coli* thymidylate synthase gene in plasmids pCG138 and pCG148 (see Table 5, of the '972 patent) and it was found to be partially effective in reducing deoxyuridine. The T4 thymidylate synthase is much more

WO 01/02580

PCT/GB00/02357

effective which is surprising in light of the fact that the *E. coli* enzyme is not thought to be controlled by any type of allosteric regulation (Neuhard and Kelln, *supra*). The *E. coli* enzyme  $K_m$  for dUMP, 4  $\mu M$  (Wahba, A. J. and M. Friedkin, *J. Biol Chem.* 237: 3794-3801), and the T4 enzyme  $K_m$  for dUMP, 2.73  $\mu M$  (Maley, F., L. LaPat-Polasko, V. Frasca and G.F. Maley, Functional domains in T4 Thymidylate Synthase as probed by site-directed mutagenesis, Chapter 29 [In] Karam, J.D. [ed] "Molecular Biology of Bacteriophage T4", American society for Microbiology, Washington, DC, 1994, pp 322-325), are similar and cannot explain the large difference in effectiveness. Whilst not being bound by theory, the inventors believe that the *E. coli thyA* has an internal transcription termination sequence derived from an upstream gene that could be effecting the expression level in plasmid clones (Bell-Penderson, D, J.L. Galloway Salvo, and M. Belfort, *J. Bacteriol.* 173: 1193-1200, 1991).

In other preferred embodiments, host cells of the present invention, particularly for use in the commercial production of pyrimidine deoxyribonucleosides e.g. thymidine comprise a transcription unit (e.g. operon) which unit comprises DNA sequences e.g. *udk* gene encoding for uridine kinase and preferably DNA sequences e.g. *dcd* gene encoding for dCTP deaminase. See e.g. Wang, L. and B. Weiss, *J. Bacteriol.* 174:5647-5653 (1992); and Neuhard, J. and L. Tarpø, *J. Bacteriol.* 175: 5742-5743.

The construct of this aspect of the invention may additionally comprise a transcription unit encoding for ribonucleotide reductase (*nrdA* and *nrdB*) and the thioredoxin (*nrdC*), or precursor forms thereof which are preferably heterologous with respect to host cell DNA and preferably derived from *E. coli* bacteriophage, particularly T "even" phages e.g. T2, T4 or T6.

Uridine kinase produces UMP and CMP from uridine and cytidine using GTP (or dGTP) as the phosphate donor. The reaction is inhibited by UTP and CTP (J. Neuhard and R.D. Kelln, Biosynthesis and Conversions of Pyrimidines, Chapter 35 [in] F.C. Neidhart et al. [eds], "*Escherichia Coli* and *Salmonella Cellular* and Molecular Biology", Second Edition, ASM press, Washington D.C). The present inventors have found that the use of uridine kinase particularly together with dCTP deaminase leads to a marked improvement in the production of thymidine by host cells incorporating these changes

together with the teachings of '972 outlined *supra*. This observation is quite unexpected since uridine kinase, on the basis of current information, has no direct role in pyrimidine *de novo* biosynthesis, moreover that its use would be beneficial in commercial processes for the production of pyrimidine deoxyribonucleosides. It is preferred that *udk* and *dcd* genes are arranged in tandem in the same operon. Further envisaged are genes that encode precursor forms of the *udk* and *dcd* gene which are processed to produce a mature form. Such processing may proceed via various intermediate forms. The *udk* and *dcd* genes may be introduced into the construct (e.g. vector) from any suitable source by methods well known to those skilled in the art for example P1 transduction, electroporation or transformation.

The enzyme uracil DNA glycosylase, encoded by the *ung* gene, is responsible for degrading DNA that has uracil incorporated in place of thymine. Where host cells of the present invention are used in the commercial production of e.g. thymidine, the internal cellular concentration of dTTP may be lowered as a result of the utilization of dTMP (a precursor of dTTP) in the production of thymidine. Accordingly, the present inventors have recognized that there is potentially a greater propensity for uracil incorporation into the host DNA which may be lethal to a wild type host due to the uracil DNA glycosylase activity causing too many single stranded breaks in the host cell DNA. Thus, host cells useful in the present invention may further display repressed (compared to the unmodified cell) or no uracil DNA glycosylase activity. This repression or absence may be achieved through various ways apparent to those skilled in the art. For example, antagonism (either total or partial) of the *ung* gene expression products is one such approach by introducing an antagonist of the functional enzyme (or precursor thereof) into the host cell. Other approaches include manipulating *ung* gene expression by e.g. modifying regulatory elements of *ung* gene expression or introducing mutations into the *ung* gene itself such that *ung* gene product expression displays little or no uracil DNA glycosylase protein and/or activity. Another approach is to delete the *ung* gene (or functionally critical parts thereof) from host cell DNA. The absence or low level of uracil DNA glycosylase activity may be a feature of the host cell without the need for further manipulation.

WO 01/02580

PCT/GB00/02357

In preferred embodiments of the present invention, each of the advances taught herein are incorporated into a host cell. The *nrd*, *td*, *udk* and *dcd* genes may be located on separate constructs but it is preferred that they are all located on the same construct e.g. vector. Thus in a particularly preferred embodiment of the present invention, a modified host cell is provided in which the cell comprises a DNA construct (e.g. vector) comprising a transcription DNA unit (e.g. operon) which unit comprises DNA sequences encoding for (preferably a T even phage, e.g. T4) a modified ribonucleotide reductase and thioredoxin in which said reductase preferably displays less sensitivity to allosteric inhibition than wild type host cell equivalent or counterpart wherein said construct further comprises:

- (a) a transcription unit (e.g. operon) encoding for (preferably heterologous with respect to host cell equivalent or counterpart) thymidylate synthase and;
  - (b) a transcription unit (e.g. operon), encoding for uridine kinase and preferably dCTP deaminase;
- and in which the host cell displays repressed or absent uracil DNA glycosylase activity.

Host cells modified according to the present invention are particularly useful in the commercial production of pyrimidine deoxynucleosides. In a particularly advantageous use of the present invention, *E. coli* host cells comprising (harboring) a plasmid modified according to the present invention (particularly in conjunction with the teachings of the '972 patent) may be used in the commercial production of thymidine. Thus, host cells modified according to the present invention may further comprise dTMPase derived from e.g. PBS1 and the mutations taught in the '972 patent, e.g. *deoA*, *tdk-1* and *udp-1*.

Generally, a fermentation method is employed which involves submerging the cells in a culture medium contained within a suitable vessel. Following culturing under appropriate conditions, produced thymidine is harvested and purified (enriched), if necessary, to pharmaceutical grade according to standard protocols. The purified thymidine may then be used in the production of medicaments, e.g. pharmaceutical compositions such as AZT.



WO 01/02580

PCT/GB00/02357

The restriction site *Xba*I was introduced at the beginning of the *nrdA* in the amplified DNA, and *Cl*aI was introduced at the 3' end of *nrdA*. The primers for PCR amplification of *nrdB* gene were:

5'-TAT ATC GAT AAA TGT AAA TTT AAG GAT TCT AAA TG-3' (Seq. id 3) and

5'-TAT GTC GAC TCC TTA AAA GTA TTT TTT AAA ACT C-3' (Seq. id 4).

The restriction site *Cl*aI was thus introduced at the beginning of the *nrdB*, and *Apa*I was introduced at the end of *nrdB* in the amplified DNA. The PCR fragments were cloned into plasmid vectors as illustrated in Fig. 1 according to techniques known to those skilled in the art. The cloned *nrdAB* genes were confirmed by enzyme activity assay. The T4 *nrdC* gene was cloned into pKC30 producing plasmid pDL51 (LeMaster, D.M., *J. Virology* 59: 759-760, 1986 ) and was supplied by D. LeMaster (Dept of Biochem., Univ. Wisconsin, Madison, WI). The gene was sub-cloned into a plasmid with *nrdAB* genes as illustrated in the Figure 1. The sources of starting materials and background information used in Figure 1 are listed in Table 1.

A synthetic transcriptional terminator was used for the construction of pCG198 and pCG301 (see Fig. 1 and Table 1). Specifically, plasmid pBC *sk*<sup>+</sup> obtained from Stratagene (La Jolla, California) was digested with restriction enzyme *Apa*I and *Asp*718 I. Synthetic DNA containing the ECOTGP transcription termination sequence (d' Aubenton Carafa *et. al.*, *J. Mol. Biol.* 216: 839-843, 1990) was then ligated replacing the original sequence

**Table 1: Genealogy of Plasmid pCG366 and pCG532 and Materials Source**

Plasmid	Gene(s)	Pro-moter	Vector origin	Marker	Derivation
PACYC177	<i>amp kan</i>		p15A	<i>amp kan</i>	Chang and Cohen (1978) <i>J. Bacteriol.</i> 134:1141-1156. ATCC 37031
pBC <i>sk+</i>	<i>Cam</i>	<i>Plac</i>	<i>ColE1</i>	<i>Cam</i>	Stratagene, 11011 North Torrey Pines Road, La Jolla, CA 92037.
pBluescript II <i>ks</i>	<i>Amp</i>	<i>Plac</i>	<i>colE1</i>	<i>Amp</i>	Stratagene, 11011 North Torrey Pines Road, La Jolla, CA 92037. ATCC 87047
pBR322	<i>amp tet</i>		<i>colE1</i>	<i>amp tet</i>	Bolivar <i>et al.</i> (1977) <i>Gene</i> 2:95-113.
pDL51	<i>T4 nrdC</i>		<i>colE1</i>	<i>Amp</i>	LeMaster, <i>J. Virol.</i> 59: 759-760, 1986
pKC30	<i>Amp</i>	$\lambda P_L$	<i>colE1</i>	<i>Amp</i>	Shimatake, and Rosenberg. (1981) <i>Nature</i> 292:128. ATCC 37286
PKTd $\Delta$ I	<i>T4 td</i>		<i>colE1</i>	<i>Amp</i>	West <i>et al.</i> (1986) <i>J.B.C.</i> 261:13446-13450.
pT7/T3/18	<i>Amp</i>	phage T7/T3	<i>colE1</i>	<i>Amp</i>	Life Technologies, INC., 9800 Medical Center Dr., Rockville, MD 20897.
pTZ18U	<i>Amp</i>		<i>colE1</i>	<i>Amp</i>	Bio-Rad, 2000 Alfred Nobel Dr., Hercules, CA 94547.
pUC18	<i>Amp</i>	<i>Plac</i>	<i>colE1</i>	<i>Amp</i>	Yanisch-Perron <i>et al.</i> (1985) <i>Gene</i> 33:103-119. ATCC 37253
pUC19	<i>Amp</i>	<i>Plac</i>	<i>colE1</i>	<i>Amp</i>	Yanisch-Perron <i>et al.</i> (1985) <i>Gene</i> 33:103-119. ATCC 37254
pCG173	$\lambda P_L$ vector	$\lambda P_L$	<i>colE1</i>	<i>Amp</i>	$\lambda P_L$ in pBluescript II <i>ks</i>
pCG196	<i>T4 nrdC</i>		<i>colE1</i>	<i>Amp</i>	<i>T4 nrdC</i> in pUC19
pCG198	<i>Cam</i>	<i>Plac</i>	<i>colE1</i>	<i>Cam</i>	Synthetic transcription terminator in pBC <i>sk+</i>
pCG301	$\lambda P_L$ vector	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	$\lambda P_L$ in pCG198
pCG308	<i>T4 nrdC</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 nrdC</i> in pCG301, LeMaster, <i>J. Virol.</i> 59: 759-760, 1986
pCG312	<i>T4 nrdA</i>	PT7	<i>colE1</i>	<i>Amp</i>	<i>T4 nrdA</i> in pT7/T3/18
pCG318	<i>NrdB</i>	PT7	<i>colE1</i>		<i>T4 nrdB</i> in pBluescript II <i>ks</i>
pCG323	<i>NrdAC</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 nrdAC</i> in pCG301
pCG326	<i>NrdABC</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 nrdABC</i> in pCG301
pCG332	<i>NrdAB</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 nrdAB</i> in pCG301
pCG334	<i>NrdAB</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 nrdAB</i> in pCG301
pCG337	<i>NrdCAB</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 nrdCAB</i> in pCG301
pCG343	<i>NrdCAB</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 nrdCAB</i> in pCG301
pCG356	<i>Td</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 td</i> in pCG301
pCG358	<i>dcd, udk</i>		<i>colE1</i> (pUC18)	<i>Amp</i>	<i>E. coli dcd, udk</i> in pUC18
pCG360	<i>NrdCAB, td</i>	$\lambda P_L$	<i>colE1</i>	<i>Cam</i>	<i>T4 nrdCAB td</i> in pCG301
pCG366	<i>NrdCAB, td</i>	$\lambda P_L$	<i>colE1</i> (pBR322)	<i>Cam Tet</i>	<i>T4 nrdCAB, td</i> in pBR322, <i>Amp<sup>S</sup></i>
pCG374	<i>dcd, udk</i>		p15A	<i>amp</i>	<i>E. coli dcd, udk</i> in pACYC177
pCG376	$\lambda P_L$ vector	$\lambda P_L$	p15A	<i>amp</i>	pACYC177 with $\lambda P_L$ & MCS of pCG301
pCG464	<i>T4 nrdA</i> <i>XbaI/HindIII</i> fragment	-	(pTZ18U) <i>colE1</i>	<i>Amp</i>	<i>XbaI/HindIII</i> fragment with <i>nrdA</i> sequence cloned into <i>XbaI/HindIII</i> sites of pTZ18U (BioRad Laboratories)



pCG492	T4 <i>nrdA</i> <i>Xba</i> I/ <i>Hind</i> III Fragment	-	(pTZ18 U) <i>colE1</i>	Amp	PCG464 with mutation in T4 <i>nrdA</i> sequence causing 79 <sup>ala</sup> → ile alteration
pCG494	<i>NrdCAB td</i>	$\lambda$ P <sub>L</sub>	<i>colE1</i> (pBR32 2)	Cam Tet	<i>Kpn</i> I/A/I fragment from pCG494 inserted into pCG366 adding 79 <sup>ala</sup> → ile mutation to <i>nrdA</i>
pCG532	<i>NrdCAB, td,</i> <i>udk, Dcd</i>	$\lambda$ P <sub>L</sub>	<i>colE1</i> (pBR32 2)	Cam	<i>udk, dcd</i> cloned into pCG494

between *Apa* I and *Asp*718 I endonuclease sites. This fragment recreated the *Apa* I recognition sequence, but destroyed the *Asp* 718 I recognition sequence in the new plasmid. The inserted DNA had the following composition

5' -CGAGC CCGCCTAATG AGCGGGCTTT TTT TT -3'  
3'-CCGGGCTCG GCGGATTAC TCGCCCGAAA AAAACATG -5'  
(shown as respective strands in seq. id 5 and 6)  
produced from two oligonucleotides.

Plasmid pCG198 was combined with pCG173 that contains the lambda P<sub>L</sub> promoter from plasmid pKC30 cloned into pBluescript II *ks* (Stratagene, La Jolla, California) as shown in Figure 1. Both pCB198 and pCG173 were digested with *Hind* III and *Asn* I then ligated to create new plasmid pCG301 containing the lambda P<sub>L</sub> promoter, multiple restriction enzyme cloning sites, and followed by the ECOTGP terminator sequence copied from the tryptophan operon leader peptide region.

T4 ribonucleotide reductase activity was measured by HPLC by a method that does not involve the use of radioisotopes and UDP substrate. The direct ribonucleotide reductase assay contained 1 mM NADPH, 1 mM DTT, 0.5 mM dATP, 0.6 mM UDP, 20 mM Tris (pH 8.0) and 5 mM MgCl<sub>2</sub>. Under these conditions the *E. coli* ribonucleotide reductase is inhibited and is not detected. The enzyme reaction (100  $\mu$ L) was stopped by the addition of 10  $\mu$ L of 50% trichloroacetic acid (TCA). After 10 minutes on ice, the samples are centrifuged in a microcentrifuge. The supernatant was extracted 4 times with diethylether to remove the TCA. Five ml of Tris buffer (1.0 M pH 8.0) was added followed by 2  $\mu$ L of 40 mg/mL rattlesnake venom (Sigma), and the sample was incubated at 37°C for 60 minutes. The samples were then heated for 3 minutes at 70°C followed by centrifugation for 5 minutes to remove precipitate. The volumes are equalized, then analysed by HPLC with an UV detector and deoxyuridine as the

standard. The column is a Spherisorb ODS-2, 5 micron, 250 mm X 4.6 mm using a 12 mM ammonium phosphate (pH 5.0) mobile phase and a flow rate of approximately 1.0 mL/minute. Results are shown in Table 2 for cells containing plasmid pCG343 demonstrating functional expression of T4 ribonucleotide reductase.

**Table 2: Ribonucleotide Reductase Activity**

<i>Strain</i>	<i>Induction condition</i>	<i>Specific Activity nmol/10 min./mg protein</i>
CMG1093	Uninduced	0
CMG1093	Induced	0
CMG1093/pCG343	Uninduced	0
CMG1093/pCG343	Induced	704.7

### **EXAMPLE 2**

#### **Derivation of Host Strain CMG2451 from Strain CMG1115**

Strain CMG1115 is fully described in McCandliss & Anderson (U. S. Patent 5,213, 972). CMG1115 was the starting point for development described herein. Strain CMG1115 was improved for thymidine productivity by selection for growth on medium containing 30 mg/L of 5-fluorouridine that yielded strain CMG2401. Strain CMG2401 was then selected for growth on medium containing 30 mg/l of 3'-azido-3'-deoxythymidine which yielded strain CMG2404. CMG2404 requires L-proline for growth due to the inherited mutation  $\Delta(lac-pro)$  from its original parent JM101. Hfr mating between CMG2404 and a Hfr stain CAG5053 (Singer, M. et al. *Microbiological Review*: 5:1-24,1989) was performed according to techniques known to those skilled in the art and yielded strain CMG2434 which is Lac<sup>+</sup>, Pro<sup>+</sup>. The *udp* (uridine phosphorylase) mutation in CMG2434 still had partial uridine phosphorylase activity that was evident based on thymine accumulation after induction of thymidine production. The *udp* mutation was reintroduced from CGSC5128 (*E. coli* Genetic Stock Center, Yale University) by phage P1 transduction according to techniques known to those skilled in the art. The *metE3079::Tn10* from strain CAG18491 was first transduced into CMG2434 to serve as a positive selection marker for transduction of *udp*. Then the *udp-1* was transduced into the *metE3079::Tn10* derivative of CMG2434 by selecting for growth without L-

WO 01/02580

PCT/GB00/02357

methionine in the defined medium. The *udp-1* derivative was named strain CMG2451. The genealogy of CMG2451 is summarized in Table 3.

**Table 3: Genealogy of *E. coli* Host Strain CMG2451**

Strain	Genotype	Derivation <sup>a</sup>
CMG1115	CMG1106 (Tn5::dTMPase kan <sup>r</sup> )	Tn5::dTMPase insertion from pCG132.
CMG2401	CMG1115 FudR <sup>r</sup>	5-fluoro-2'-deoxyuridine resistance.
CMG2404	CMG2401 AZT <sup>r</sup>	3'-Azido-3'-deoxythymidine resistance.
CMG2434	CMG2404 Lac <sup>+</sup> Pro <sup>+</sup>	Repair $\Delta(lac-proAB)$ by conjugation with CAG5053 <sup>b</sup> .
CMG2448	CMG2434 <i>metE</i> 3079::Tn10	<i>metE</i> 3079::Tn10 from CAG18491 <sup>b</sup> .
CMG2451	CMG2448 <i>udp-1 metE<sup>+</sup> tet<sup>r</sup></i>	<i>udp-1</i> from CGSC5128 <sup>c</sup> and replaced <i>metE</i> 3079::Tn10.

<sup>a</sup>. Mutations were at times introduced into *E. coli* strains by phage P1 transduction. If the mutation has a selective marker, direct P1 transduction was used. If the mutation has no selective marker, P1 cotransduction with the nearby Tn10 insertion was used.

<sup>b</sup> Singer, M., et al. *Microbiological Review*: 53:1-24, 1989.

<sup>c</sup>. All CGSC strains can be obtained for *E. coli* Genetic Stock Center, Yale University, P.O. Box 208104, New Haven, CT 06520-8104.

### EXAMPLE 3

#### **Cloning and expression of T4 *td* gene into thymidine production plasmid**

The T4 *td* gene was cloned into pKTd $\Delta$ I by West et al. (*J. Biol. Chem* 261:13446-13450, 1986) without the 1017-base pair intron. The *td* gene was sub-cloned into pCG301 using two oligonucleotides as linkers with the following sequences:

5'-GAT CCG GAG GAT AAA TGA AAC AAT ACC AAG ATT TAA T-3' (seq. id 7) and;  
5'-TAA ATC TTG GTA TTG TTT CAT TTA TCC TCC G-3' (seq.id 8).

The result plasmid was pCG356. The *td* gene from pCG356 was then sub-cloned into pCG343 to create pCG360 (Fig. 1). The tetracycline resistance gene and plasmid replication origin from pBR322 (Bolivar, F. et al., *Gene* 2: 95-113, 1977) was sub-cloned into pCG360 and formed pCG366. The thymidylate synthase activity was

WO 01/02580

PCT/GB00/02357

measured by spectrophotometric method of Wahba and Friedkin (*J. Biol.Chem.* 236: PC11-PC12, 1961 ). The results are shown in Table 4.

**Table 4: Thymidylate Synthase Activity**

Strain	Specific Activity $\Delta OD_{340}/mg \text{ protein}/min$
CMG1093	-0.72
CMG1093/ pKTdDI	3.24
CMG1093/pCG356	3.45
CMG1093/pCG360	1.43

#### EXAMPLE 4

##### **Shake flask data showing the value of T4 *td* gene on thymidine production and deoxyuridine reduction**

The shake flask fermentation was used to evaluate the thymidine productivity of different *E. coli* recombinants. The shake flask fermentation broth and methods used here, and in other examples, is described in Example 6 below. In this case 20 mL volume per flask was inoculated with 2 ml of seed culture and incubated in a 30°C shaker. At about 10 OD 600nm, the flasks were transferred into a 37°C shaker for 30 minutes to mildly induce the  $\lambda$  P<sub>L</sub> promoter. Then the flasks were transferred into a 35°C shaker to continue the fermentation. Glucose was fed during fermentation as needed, and pH was adjusted to about 7 with ammonia according to the color of phenol red. The thymidine concentration was measured by HPLC using a Spherisorb ODS-2, 5 micron, 250-mm X 4.6 mm column, and a 25 mM ammonium phosphate (pH 3.3) mobile phase with a flow rate of approximately 1.5 ml/minute.

The strain CMG2451/pCG366 (T4 *nrdCAB*, *td*) was compared with CMG2451/pCG343 (T4 *nrdCAB*) in shake flask fermentation. Table 5 shows that deoxyuridine concentration was reduced, and converted to thymidine in strain CMG2451/pCG366 due to the T4 *td* gene.

**Table 5: Thymidine Production-Effect of Thymidylate Synthase at 66 hours.**

<i>Stain</i>	<i>Thymidine (mg/l)</i>	<i>Deoxyuridine (mg/l)</i>	<i>% Deoxyuridine</i>
CMG2451/pCG343	1198	1728	144.2
CMG2451/pCG366	3327	206	6.2

### EXAMPLE 5

#### Construction of the <sup>79</sup>Ala to Ile T4 ribonucleotide reductase mutant

The mutation was introduced by using site-directed mutagenesis based on a method described by Kunkel (Kunkel, T.A., *Proc. Natl. Acad. Sci. USA*, 82: 488-492, 1985). All materials for mutant construction including pTZ18U phagemid DNA, M13KO7 helper phage, bacterial strains *E. coli* CJ236 and MV1190, T7 DNA polymerase and T4 DNA ligase were provided in the Muta-Gene *in vitro* mutagenesis kit from Bio-Rad Laboratories (Hercules, CA). At first, the *Xba*I/*Hind*III DNA fragment containing the *nrdA* gene of T4 bacteriophage was isolated from plasmid pCG312 (ChemGen Corp., Table 1 above). The *Xba*I/*Hind*III DNA fragment was cloned into the *Xba*I/*Hind*III sites of the pTZ18U phagemid vector using standard protocols (Sambrook, J., Fritsch, E.F. and Maniatis, T., *Molecular Cloning*, Cold Spring Harbor Laboratory, New York, 1989). The insert-carrying phagemid pCG464 was introduced into *E. coli* CJ236. This strain is deficient for dUTPase (*dut*) and uracil-N-glycosylase (*ung*) which results in an occasional substitution of uracil for thymine in newly synthesized DNA. Single stranded DNA of pCG464 containing uracil was isolated from CJ236 according to Bio-Rad Laboratories Instructional Manual. This DNA (0.2 pMole) was annealed with 6 pMole of phosphorylated primer

5'-AGC AAA CAT TAA ACA GCG TGC AATTAC ATA TTG ATA ATC AGG TTC-3'

(sequence id 9) containing the sequence of the desired mutation (underlined) coding for Ile instead of original <sup>79</sup>Ala.

Complementary strand DNA was synthesized by using T7 DNA polymerase as described in the Bio-Rad protocol. The reaction products were transformed into *E. coli* MV1190 containing a wild type uracil-N-glycosylase, which degrades the uracil-containing parental strand, thus enriching for the mutant strand. Direct DNA sequencing

WO 01/02580

PCT/GB00/02357

using the Silver Sequence DNA Sequencing System from Promega Corp. (Madison, WI) identified plasmids containing the desired mutation. It appeared that all four analysed transformants contained plasmids with the mutation. One of them was designated pCG492 and used for further experiments. To check if the mutation affects thymidine synthesis it was introduced into production plasmid pCG366 (ChemGen Corp., Table 1 and Figure 1). For this, the *KpnI/AflII* DNA fragment of pCG366 containing 5'-part of the *nrda* gene was replaced with *KpnI/AflII* fragment from pCG492 that contains the mutation. The new plasmid pCG494 was introduced into production strain CMG2451 (ChemGen Corp., Example 2 and Table 4 above). The effect of T4 *nrda* mutation was evaluated by comparison of thymidine production by CMG2451 (pCG494) and CMG2451 (pCG366) as shown in Example 6 below.

#### **EXAMPLE 6**

##### **Shake flask fermentation for Thymidine production using CMG2451 (pCG366) and CMG2451 (pCG494)**

The 250-mL baffles flasks containing 25 mL of production medium were inoculated with 2 mL of a freshly grown seed culture in LB broth with appropriate antibiotic added. The cultures were grown in a 30°C shaker at 250 rpm. When the OD<sub>600</sub> reached about 5, the flasks were transferred into a 37°C shaker for 30 min. Then the flasks were transferred into a 35°C shaker to continue the fermentation. The production medium has the following composition (g/L): Ardamine YEP-S (Red Star Yeast & Products, Milwaukee, WI) - 10; CaCO<sub>3</sub> - 10; MgSO<sub>4</sub> - 0.4; phenol red - 0.24; PP90BT (DMV International, Fraser, N.Y.) - 4.5; sorbitol - 20; chloramphenicol - 0.03; trace elements (1000X) - 1 mL/L. The trace elements (1000X) formulation is the following (g/L): boric acid - 0.05; calcium chloride - 20; cobalt sulfate - 0.05; copper sulfate - 0.01; ferrous sulfate - 20; ferric chloride - 20; manganese sulfate - 0.5; sodium molybdenate - 0.1; and zinc sulfate - 0.1. At the time of induction 10 grams per liter of Ardamine YEP-S was added. Glucose was fed during the fermentation on an average of every two hours (2.5 g/L). The pH in the flasks was maintained at approximately 7.0 through the addition of 4N NH<sub>4</sub>OH as judged by the color of the phenol red indicator dye. The OD<sub>600</sub> was read after sample dilution 1:10 into 10 mM H<sub>2</sub>SO<sub>4</sub> to dissolve salts in the medium. Thymidine concentration was measured by reverse phase C-18 HPLC with an Alltech Spherisorb

WO 01/02580

PCT/GB00/02357

ODS-2 column and Shimadzu spectrophotometric detector at 260 nm. The mobile phase was a 25 mM  $\text{NH}_4\text{H}_2\text{PO}_4$  (pH 3.3) in water at the constant rate of 1.5 mL/min.

The results of thymidine production by CMG2451 (pCG366) and CMG2451 (pCG494) after 2, 17 and 25 hours after induction are presented in Table 6. There were two repeats of the flasks in this experiment and variability between duplicate flasks did not exceed 15%. The results show that the T4 *nrdA* mutant performed better than the wild type *nrdA* strain. This was confirmed in several independent shake flask experiments with the same bacterial strains.

**Table 6: Thymidine production by CMG2451 (pCG366) and CMG2451 (pCG494)**

Strain	2 hours	17 hours	25 hours
Specific activity (mg/L/OD)			
CMG2451 (pCG366)	6.1	32.5	43.1
CMG2452 (pCG494)	8.5	36.1	51.4

### EXAMPLE 7

#### Shake Flask Fermentation to Demonstrate the Effect of the *E. coli* *udk* Gene on Thymidine Production.

The *dcd* gene or *dcd* *udk* operon were cloned into the pACYC177 vector. This vector with the p15a origin of replication is different from *colE1* based plasmids such as pCG366 and thus is compatible and can be maintained in the same host with *colE1* based plasmids. The details of the plasmid constructions resulting in pCG374 (*udk dcd*) or pCG375 (*dcd*) are shown in Figure 2. The genes on these plasmids are expressed from the native *E. coli* promoter of the *udk dcd* operon.

Using selection for ampicillin resistance, plasmids pCG374 and pCG375 were introduced into CMG2451 (pCG366) to test the effect on thymidine production in the shake flask fermentation method described in Example 6. The results at several time points are shown in Table 7. Although the specific activity (thymidine per OD of cells) is

WO 01/02580

PCT/GB00/02357

similar both with and without the *udk* gene, the cells with the *udk* gene on the second plasmid pCG374 grew to a higher cell density and produced significantly more thymidine (5.8 g/L compared to 3.0 g/L for the strain with the *dcd* only second plasmid). This result was not anticipated, as it is not clear why uridine phosphorylase could have this effect.

Based on this data and other information, the *udk* gene was chosen for introduction along with the *E. coli dcd* gene in the construction of plasmid pCG532 (see Example 8 and Figure 3).

**Table 7: Comparison of the effect of second plasmids with *dcd* or *dcd udk* on thymidine production in the CMG2451 (pCG366) background.**

Time (hour)	O.D.600		Thymidine (mg/l)		Specific Thymidine (mg/l/OD)	
Base Strain And Plasmid	CMG 241 (pCG 366)	CMG 2451 (pCG 366)	CMG2451 (pCG366)	CMG2451 (pCG366)	CMG2451 (pCG366)	CMG2451 (pCG366)
Second pACYC 177 based plasmid	PCG 374 with <i>dcd udk</i>	pCG375 with <i>dcd</i>	pCG374 with <i>dcd udk</i>	pCG375 with <i>dcd</i>	pCG374 with <i>dcd udk</i>	PCG375 with <i>dcd</i>
8 hr	29.112	29.94	889	884	30.5	29.5
18 hr	40.206	32.946	1721	1608	42.8	48.8
42 hr	48.348	33.96	4020	2740	83.1	80.7
66 hr	57.498	28.614	5804	3012	100.9	105.3

### EXAMPLE 8

#### Cloning of the *E. coli udk dcd* operon into production plasmid pCG494.

The *udk* and the *dcd* genes of *E. coli* encode pyrimidine ribonucleotide kinase and dCTP deaminase, respectively. Both genes were mapped to a 3.4 kb *Bam*HI/*Pst*I DNA fragment of lambda phage 355 of the Kohara genomic library (Kohara, Y., Akiyama, K. and Isono, K., *Cell* 50, 495-508, 1987). It appears that *udk* is located upstream of *dcd* and transcribed in the same direction as *dcd* (Neuhard, J. and Tarpo, L., *J. Bacteriol.* 175: 5742-5743, 1993). The genes were cloned into production plasmid by two steps.



WO 01/02580

PCT/GB00/02357

At first, a 3.4 kb *Bam*HI/*Pst*I DNA fragment from lambda 355 was cloned into the multiple cloning site of plasmid pUC18 (Yamisch-Perron, C., Vieira, J. and Messing, J., *Gene* 33: 103-109,1985) (plasmid pCG358). Then, the fragment was excised from the polylinker region of pCG358 with *Bam*HI and *Sph*I and cloned in place of a 0.7 kb *Bgl*II/*Sph*I fragment (containing a portion of the tetracycline resistance gene) of production plasmid pCG494. The final 14.7 kb plasmid pCG532 contains *col*E1 compatibility group origin of replication, the chloramphenicol resistance gene, the *udk* and the *dcd* genes and the T4 bacteriophage *nrdCAB* (encode thioredoxin and two subunits of ribonucleotide reductase, respectively) and the T4 *td* (encodes thymidylate synthase) genes under control of the P<sub>L</sub> promoter of bacteriophage lambda. The T4 *nrdA* gene of pCG532 was previously changed by site-directed mutagenesis (<sup>79</sup>Ala to Ile).

A synthetic transcriptional terminator is located downstream of the *td* gene to prevent transcriptional readthrough into the replication region. The genetic map of plasmid pCG532 is illustrated in Figure 3.

### EXAMPLE 9

#### Shake flask fermentation of thymidine by CMG2451 (pCG494) and CMG2451(pCG532)

Plasmid pCG532 containing the *udk* and the *dcd* genes of *E. coli* was introduced into production strain CMG2451. New strain CMG2451 (pCG532) was tested together with parent strain CMG2451 (pCG494) in shake flask experiments to compare thymidine production. The results of two independent experiments are shown in Table 8. The first experiment was performed as described above and samples were taken at 17 hours after induction. In the second experiment cells were induced at higher OD (about 9) and samples were taken at 3 hours after induction for analysis. In both cases the strain containing the cloned *udk* and *dcd* genes performed better than the parent strain.

**Table 8: Thymidine fermentation by thymidine by CMG2451 (pCG494) and CMG2451 (pCG532)**

<b>Experiment 1</b>			
Strain	O.D 600	TdR (mg/L)	Specific Activity(mg/L/OD)
CMG2451(pCG494)	16.5	599	36.3
CMG2451(pCG532)	17.5	867	49.5
<b>Experiment 2</b>			
Strain	O.D 600	TdR (mg/L)	Specific Activity(mg/L/OD)
CMG2451(pCG494)	8.5	149	17.5
CMG2451(pCG532)	8.4	192	22.8

**Example 10: Addition of *ung* mutation and its effect on thymidine production in Shake flask fermentation**

An uracil DNA glycosylase negative strain was constructed by introducing an *ung*::Tn10 (Varshney, U., et al., *J. Biol. Chem.* 263:7776-7784, 1988) mutation into host CMG2451 using P1 transduction as described above, and was named CMG2492. Plasmid pCG366 was introduced into CMG2492. A comparison experiment between *Ung*<sup>-</sup> and *Ung*<sup>+</sup> strain for thymidine synthesis in shake flasks is shown in Figure 4 using the flask method described in Example 6. Cells were grown in 250ml flasks at 30°C, and thymidine synthesis was induced by shifting temperature to 37°C for 30 min, then, shifting to 35°C. The *Ung*<sup>-</sup> host without uracil DNA glycosylase kept growing longer and made 30% more thymidine.

**EXAMPLE 11**

**Thymidine Production in a 30 Liter Fermentor with strain CMG2451 (pCG532)**

The following conditions were used to produce thymidine in a 30-Liter fermentor (B. Braun Biotec Biostat C) with strain CMG2451/532. The seed culture (500 mL) was grown in a 4 Liter baffle shake flask in LB medium (5 g/L Difco yeast extract, 10 g/L Difco tryptone, 5 g/L NaCl) with 30 mg/L chloramphenicol and 25 mg/L kanamycin at 30 °C until 2.37 OD 600 nm was reached with a final pH of 6.69.

WO 01/02580

PCT/GB00/02357

The initial batch in the fermentor (12 Liters) containing the composition listed in Table 9 was sterilized at 121°C for 55 minutes. After cooling a separately autoclaved solution (500 mL) was added to adjust the batch to 20 g/L sorbitol and 3.0 g/L  $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ . Also added before inoculation was a sterile filtered solution (200 mL) designed to to  
 5 adjust the initial batch to 30 mg/L chloramphenicol, 25 mg/L kanamycin, 1 mg/L d-biotin, 10 mg/L thiamine and 10 mg/L nicotinic acid.

Three feed solutions were prepared: a) Cerelese 2001 (dextrose monohydrate) 562 g/L with 2 mg/L biotin, 20 mg/L thiamine, 20 mg/L nicotinic acid, and 30 mg/L  
 10 chloramphenicol; b) sorbitol 717 g/L with with 4 mg/L biotin, 40 mg/L thiamine, 40 mg/L nicotinic acid, and 60 mg/L chloramphenicol; and c) crude nitrogen mixture containing 360 g/L Amberex 695 AG (Red Star Yeast & Products, Milwaukee, Wisc.), 6 g/L PP90M (DMV International, Fraser, NY) with 1X trace elements and 0.1 mL/L Mazu  
 DF10PMOD11 (BASF). The feed solutions were sterilized for 40 to 50 minutes under 18  
 15 PSI steam pressure.

**Table 9 : Initial batch composition in 30 L fermentor**

<i>Component</i>	<i>Concentration (g/L or mL/L)</i>
Sodium hexametaphosphate	4
$\text{KH}_2\text{PO}_4$	4
$(\text{NH}_4)_2\text{HPO}_4$	4
$\text{CaCl}_2$	0.4
Citric acid	0.5
Amberex 695 (Red Star Yeast & Products, Milwaukee, WI)	20
PP90M (DMV International, Fraser, NY)	15
Tryptone (Difco, Detroit, MI)	5
1000X Trace elements (see Example 7)	1 mL/L
$\text{CaCO}_3$	5
Glycine	1.83
Mazu DF10PMOD11 Defoamer (BASF Speciality Products, Gurnee, IL)	0.2 mL/L

The operating conditions were as follows: initial temperature 31°C; RPM 600; air flow 3 LPM; pH 6.8; pressure 0 Bar. The dissolved oxygen was controlled at 25% saturation

WO 01/02580

PCT/GB00/02357

using air flow rate control loop. The RPM was increased to 750 RPM at 8 hours, 850 RPM at 9 hours and 950 RPM at the time of temperature shift from 31 to 35.5°C at 9.2 hours when the culture reached 37.8 OD. Beginning at 9.7 hours back pressure was applied up to a maximum of 0.6 Bar to aid in oxygen transfer into the culture. The rate of temperature shift for induction of thymidine synthesis was 0.2°C per minute.

After the cell mass reached 20 OD at 600 nm (read after dilution in 50 mM H<sub>2</sub>SO<sub>4</sub> to dissolve salts), 85 mL batch feeds of the sorbitol feed solution (b) were made for each 5 OD increase in cell mass. At 16.7 hours sorbitol batch feeding was stopped and dextrose monohydrate (glucose) feed (a) was started under the control of the DO PID control loop of the B. Braun fermentor with a set point of 25%. Simply stated, when dissolved oxygen was below 25% sugar feed was off, and when DO was greater than 25% the sugar feed was set to on. The protocol self regulates the glucose concentration keeping the concentration low, but does not allow the culture to starve for glucose for a very prolonged length of time. Crude nitrogen feeds (500 mL) were made at 11 hours, 16.2 hours, 21.2 hours, 28.2 hours, 33.2 hours and 40.5 hours.

The cell mass, thymidine and deoxyuridine accumulation during the fermentation are shown in Figure 5.

## EXAMPLE 12

### Purification of Thymidine from Fermentation broth.

Dowex Optipore L-285 (The Dow Chemical Company) was suspended in deionized water and packed into a 48mm diameter glass column making a bed volume of about 500 mL. The column was washed with 500 mL of 5% NaOH, then washed with deionized water to until the effluent was pH 7.0.

500 mL Fermentation broth with thymidine (TdR) concentration of 4.890 g/L and deoxyuridine (UdR) concentration of 1.040 g/L was loaded onto the column. The column was washed with two bed volumes of deionized water. TdR and UdR were eluted by two bed volumes of 5% reagent alcohol (Ethanol 90.5%, Methanol 4.5%, Isopropyl alcohol 5%) followed by two bed volumes of 10% reagent alcohol and two

WO 01/02580

PCT/GB00/02357

bed volumes of 15% reagent alcohol. The TdR and UdR in the 25 mL fractions is shown in Fig. 6. The column was regenerated by washing with 5% NaOH, then with deionized water to pH7.0 and the procedure was repeated. The fractions 91-160 were pooled together from the two separate runs and dried using a rotary vacuum evaporator. Thymidine was re-dissolved in a minimal amount of hot water and crystallized at 4°C. Then the crystals were re-crystallized two times and dried in a 55°C oven for 15 hours. A total of 979.8 mg of crystalline thymidine was obtained with a purity of greater than 99% that should be suitable for use as a pharmaceutical intermediate.

The side fractions containing both deoxyuridine and thymidine were pooled with mother liquors and reduced by a rotary vacuum evaporator to a final volume of 1200 mL with reduced alcohol. The total amount of TdR in this 1200 mL solution was 3571 mg. The total recovery (including the 3571 mg of TdR side fractions) was 93.1%.

Claims

1. A DNA construct comprising a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene or a uridine kinase gene and/or a dCTP deaminase gene.
2. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a ribonucleotide reductase gene and a thioredoxin gene.
3. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a uridine kinase gene and/or a dCTP deaminase gene.
4. A DNA construct as claimed in claim 1 comprising a transcriptional unit which comprises a uridine kinase gene and a dCTP deaminase gene.
5. A DNA construct according to any of claims 1-4 wherein the construct is an extrachromosomal vector.
6. A DNA construct according to claim 5 wherein the vector is a plasmid, virus, transposon, minichromosome or phage.
7. A DNA construct according to claim 6 wherein the vector is a plasmid, wherein the ribonucleotide reductase gene is a *nrdA* gene and/or a *nrdB* gene, and wherein the thioredoxin gene is a *nrdC* gene and said genes are arranged in an operon.
8. A DNA construct according to claim 7 wherein the transcriptional unit comprises a *nrdA* gene and a *nrdB* gene.
9. A DNA construct according to claim 8 in which the *nrdA* and *nrdB* genes are located downstream of the *nrdC* gene.
10. A DNA construct according to claim 8 wherein the *nrdC* gene is upstream of the *nrdA* gene and *nrdA* gene is upstream of the *nrdB* gene.

11. A DNA construct according to claim 6, further comprising a regulatory element.
12. A DNA construct according to claim 11, wherein the regulatory element is selected  
5 from the group consisting of a promoter, an operator, a termination sequence, an initiation sequence and a ribosome binding site.
13. A DNA construct according to claim 12 wherein the promoter is the lambda P<sub>L</sub> promoter or a derivative therefrom.
- 10 14. A DNA construct according to claim 12 wherein the termination sequence is a synthetic terminator.
- 15 15. A DNA construct according to claim 7 wherein the *nrdA* gene is modified such that ribonucleotide reductase encoded by the unit is less sensitive to allosteric inhibition than the ribonucleotide reductase encoded by the unit comprising an unmodified *nrdA* gene.
- 20 16. A DNA construct according to claim 15, wherein the *nrdA* gene is modified at a dTTP binding site.
17. A DNA construct according to claim 15 wherein the *nrdA* gene is modified so as to encode an Ala-to-Ile change at position 79 in the *nrdA* expression product.
- 25 18. A DNA construct according to claim 7 wherein the *nrdA*, *nrdB* and *nrdC* genes are derived from a T phage.
19. A DNA construct according to claim 18 wherein the T phage is a T even phage.
- 30 20. A DNA construct according to claim 19 wherein the T even phage is a T4 phage.
21. A DNA construct according to any of claims 1-4 wherein the construct further comprises a thymidylate synthase gene.

22. A DNA construct according to claim 21 wherein the thymidylate synthase gene is the *td* gene.
- 5 23. A DNA construct according to claim 22 wherein *td* gene is located on the same vector as the *nrdA*, *nrdB* and *nrdC* genes.
24. A DNA construct according to claim 22 wherein the *td* gene is located in the same operon as the *nrdA*, *nrdB* and *nrdC* genes.
- 10 25. A DNA construct according to claim 24 wherein the *td* gene is located downstream (in terms of reading frame) from the *nrdA*, *nrdB* and *nrdC* genes.
26. A DNA construct according to claim 22 wherein the *td* gene is derived from a T phage.
- 15 27. A DNA construct according to claim 26 wherein the *td* gene is derived from a T even phage.
28. A DNA construct according to claim 27 wherein the *td* gene is derived from T4 phage.
- 20 29. A DNA construct according to claim 2 further comprising a uridine kinase gene and/or a dCTP deaminase gene.
- 25 30. A DNA construct according to claim 29, wherein the DNA construct comprises both a uridine kinase gene and a dCTP deaminase gene.
31. A DNA construct according to claim 1,3 or 4 wherein the uridine kinase gene is a *udk* gene.
- 30 32. A DNA construct according to claim 1,3 or 4 wherein the dCTP deaminase gene is a *dcd* gene





WO 01/02580

PCT/GB00/02357

- (b) a transcription unit located on said DNA construct, comprising a uridine kinase gene;
- (c) a transcription unit located on said DNA construct, comprising a dCTP deaminase gene; and
- (d) repressed or absent uracil DNA glycosylase activity.

41. A modified host cell according to claim 40, wherein the ribonucleotide reductase gene is modified at a dTTP binding site.

42. A modified host cell according to claim 41, wherein the ribonucleotide reductase is the *nrdA* gene and the modification is an Ala-to-Ile change at position 79 in the *nrdA* expression product.

43. A modified host cell according to claim 38, wherein the DNA construct comprises both the uridine kinase gene and the dCTP deaminase gene.

44. A modified host cell according to claim 40, wherein the cell comprises each one of the features of (a) to (d).

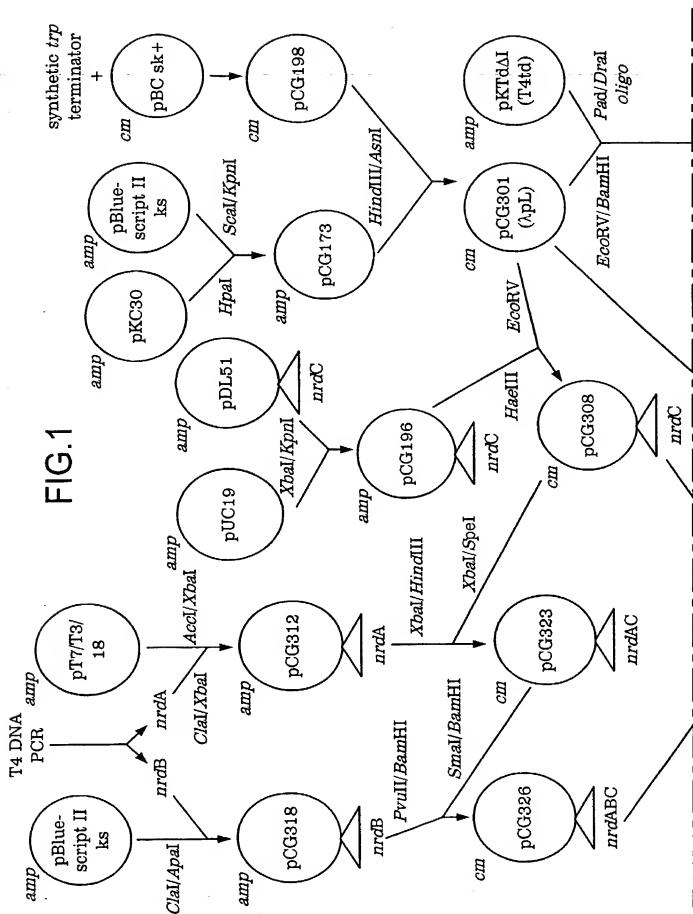
45. A process for the production of pyrimidine deoxyribonucleosides comprising culturing a host cell according to any of claims 33-44.

46. A process according to claim 45 wherein the deoxyribonucleoside is thymidine.

47. A process for the manufacture of azidothymidine comprising culturing a host cell according to any of claims 33-44.

48. A culture medium comprising a host cell according to any of claims 33-44.

1/7



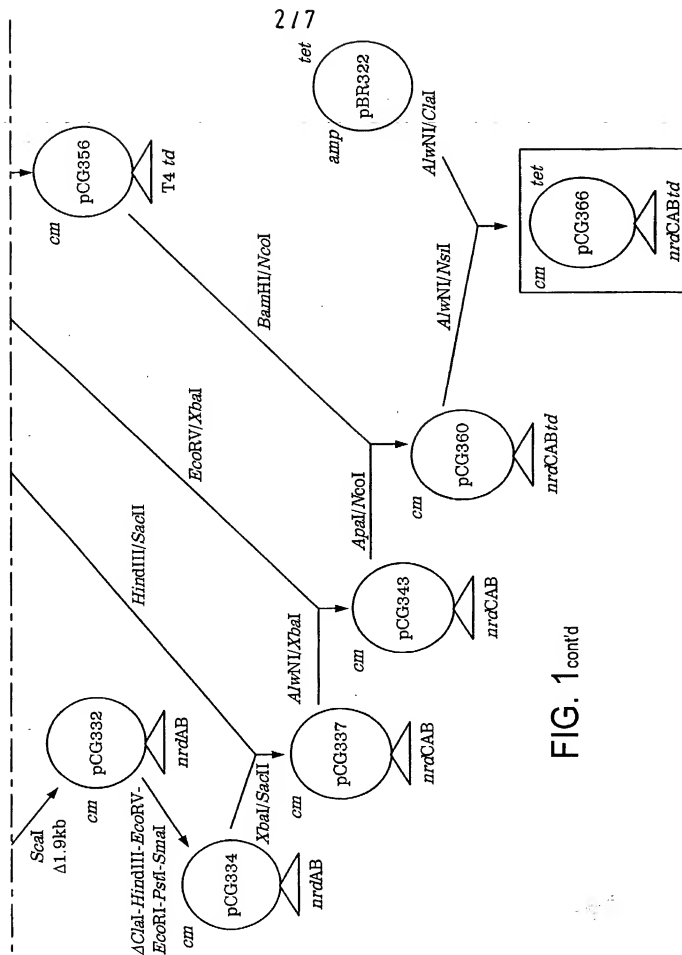


FIG. 1 cont'd

3/7

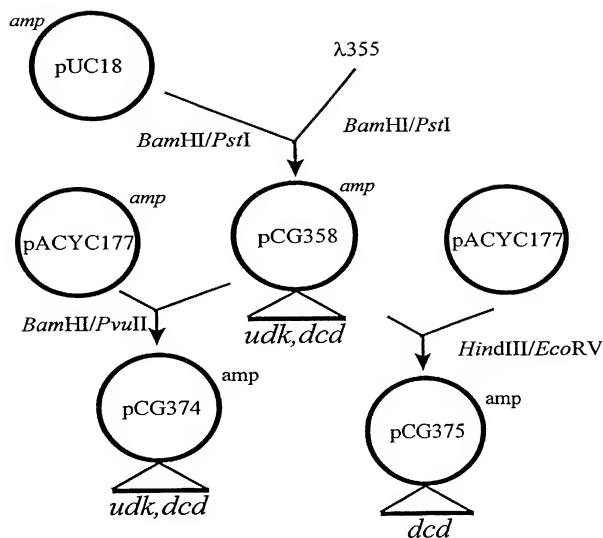
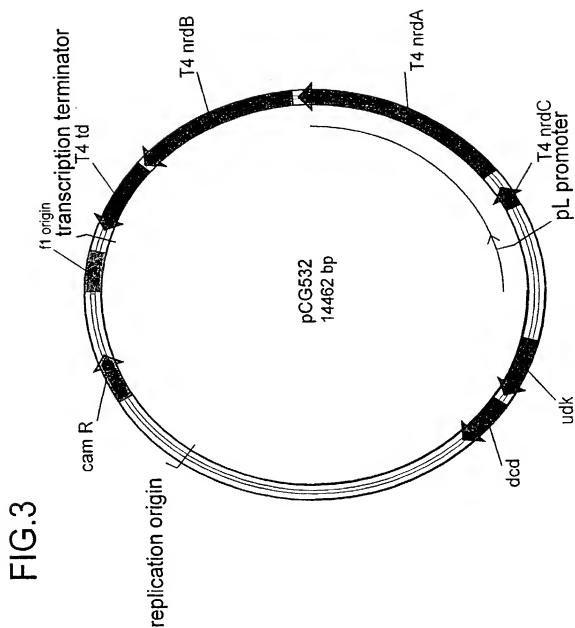
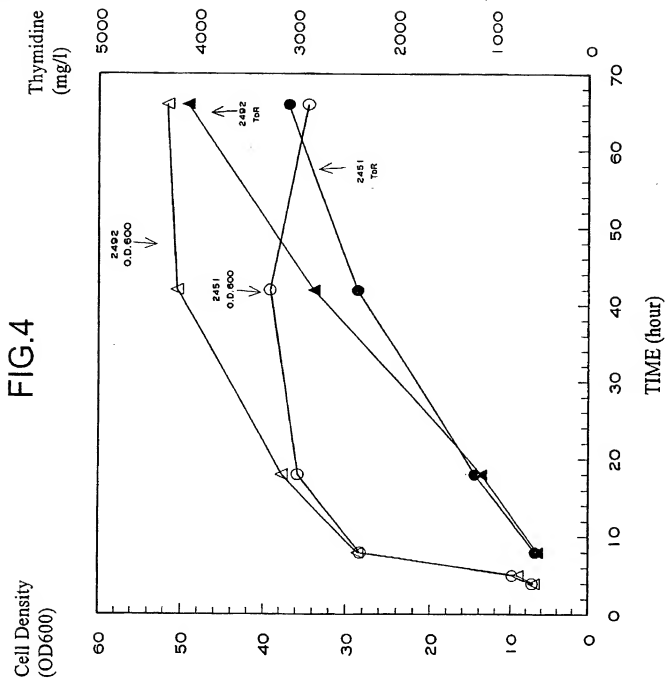


FIG.2

4/7

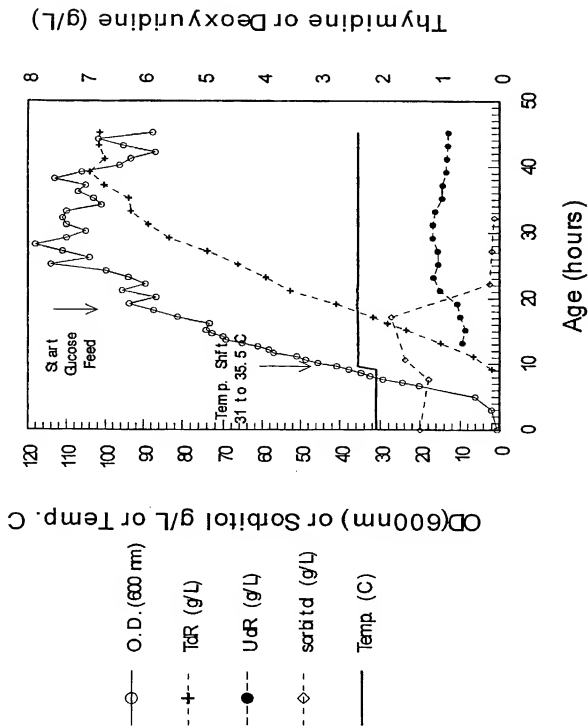


5/7



6/7

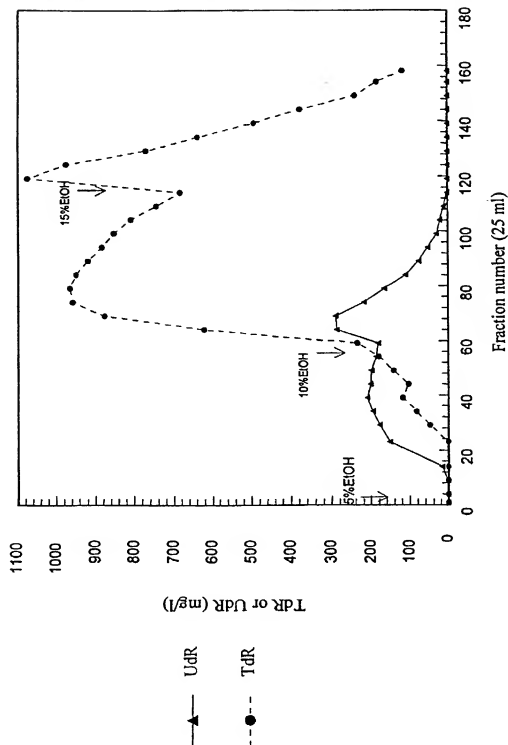
# Thymidine Production 30 Liter Fermentor Example





7/7

FIG.6

TdR purification by Optipore L285  
(500ml column)

**COMBINED DECLARATION FOR UTILITY OR DESIGN PATENT APPLICATION WITH POWER OF ATTORNEY**ATTORNEY'S DOCKET  
PX3813USWFirst Names Inventor:  
**ANDERSON****Complete if known:**App No.:  
10/018,878Filing Date  
December 19, 2001

Group Art Unit:

( ) Declaration submitted with initial filing or

(X) Declaration submitted after initial filing (surcharge required 37CFR1.16(c))

As below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

**VECTORS, CELLS AND PROCESSES FOR PYRIMIDINE DEOXYRIBONUCLEOSIDES PRODUCTION**

the specification of which (check only one item below):

[ ] is attached hereto.

OR

[ X ] was filed on \_\_\_\_\_ as United States application Serial No. 10/018,878 or PCT International

Application Number PCT/GB00/02357 filed 30 June 2000 and was amended on (MM/DD/YYYY)  
(if applicable)

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR §1.56.

I hereby claim foreign priority benefits under 35, U.S.C. §119 (a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or of any PCT international application having a filing date before that of the application on which priority is claimed:

Prior Foreign Application Number (s)	Country	Foreign Filing Date (MM/DD/YYYY)	PRIORITY CLAIMED
1.			
2.			
3.			
4.			
5.			

I hereby claim the benefit under Title 35, United States Code §119(c) of any United States provisional application(s) listed below:

Application No.	Filing Date (MM/DD/YYYY)	
1. 60/141,827	07/01/1999	X
2.		
3.		
4.		
5.		

<b>COMBINED DECLARATION FOR UTILITY or DESIGN PATENT APPLICATION WITH POWER OF ATTORNEY</b> Continued	ATTORNEY'S DOCKET NUMBER <b>PX3813USW</b>
---	--

I hereby claim the benefit under 35, U.S.C. §120 of any United States application or §365(c) of any PCT international application designating the United States of America that is listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. §112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 C.F.R. §1.56 which became available between the filing date of the prior application(s) and the national or PCT international filing date of this application.

<b>PRIOR U.S. PARENT APPLICATION or PCT PARENT APPLICATION</b>				
--	--	--	--	--

U.S. Parent Application or PCT Parent Number	Parent Filing Date (MM/DD/YYYY)	STATUS (Check one)		
09/345,492	07/01/1999	PATENTED	PENDING	ABANDONED
			X	

**POWER OF ATTORNEY:** As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the U.S. Patent and Trademark Office connected therewith. (List name and registration number)



23347

PATENT TRADEMARK OFFICE

Send Correspondence to:



23347

PATENT TRADEMARK OFFICE

Direct Telephone Calls to

Frank P. Grassler  
919-483-2482

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1-00	2	FULL NAME OF INVENTOR <b>ANDERSON</b> INVENTOR'S SIGNATURE <i>[Signature]</i>	FAMILY NAME <b>ANDERSON</b> CITY <b>Rockville</b>	FIRST GIVEN NAME <b>David</b> STATE OR FOREIGN COUNTRY <b>Maryland MD</b>	SECOND GIVEN NAME/INITIAL <b>Martin</b> DATE: <b>2/27/02</b>
0	1	RESIDENCE & CITIZENSHIP <b>Rockville</b> POST OFFICE ADDRESS <b>GlaxoSmithKline Five Moore Drive, PO Box 13398</b>	CITY <b>Rockville</b> POST OFFICE ADDRESS <b>GlaxoSmithKline Five Moore Drive, PO Box 13398</b>	CITY <b>Research Triangle Park</b>	COUNTRY OF CITIZENSHIP <b>US</b> STATE & ZIP CODE/COUNTRY <b>NC 27709 US</b>
2-00	2	FULL NAME OF INVENTOR <b>LIU</b> INVENTOR'S SIGNATURE <i>[Signature]</i>	FAMILY NAME <b>LIU</b> CITY <b>Rockville</b>	FIRST GIVEN NAME <b>Lin</b> STATE OR FOREIGN COUNTRY <b>Maryland MD</b>	SECOND GIVEN NAME/INITIAL <b>Lin</b> DATE: <b>2/27/02</b>
0	2	RESIDENCE & CITIZENSHIP <b>Rockville</b> POST OFFICE ADDRESS <b>GlaxoSmithKline, Inc. Five Moore Drive, PO Box 13398</b>	CITY <b>Rockville</b> POST OFFICE ADDRESS <b>GlaxoSmithKline, Inc. Five Moore Drive, PO Box 13398</b>	CITY <b>Research Triangle Park</b>	COUNTRY OF CITIZENSHIP <b>US</b> STATE & ZIP CODE/COUNTRY <b>NC 27709 US</b>
3-00	2	FULL NAME OF INVENTOR <b>PODKOYVROV</b> INVENTOR'S SIGNATURE <i>[Signature]</i>	FAMILY NAME <b>PODKOYVROV</b> CITY <b>Gaithersburg</b>	FIRST GIVEN NAME <b>Sergey</b> STATE OR FOREIGN COUNTRY <b>Maryland MD</b>	SECOND GIVEN NAME/INITIAL <b>Sergey</b> DATE: <b>2/27/02</b>
0	3	RESIDENCE & CITIZENSHIP <b>Gaithersburg</b> POST OFFICE ADDRESS <b>GlaxoSmithKline Five Moore Drive, PO Box 13398</b>	CITY <b>Gaithersburg</b> POST OFFICE ADDRESS <b>GlaxoSmithKline Five Moore Drive, PO Box 13398</b>	CITY <b>Research Triangle Park</b>	COUNTRY OF CITIZENSHIP <b>US</b> STATE & ZIP CODE/COUNTRY <b>NC 27709 US</b>

2	FULL NAME OF INVENTOR	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME/INITIAL
4-00	INVENTOR'S SIGNATURE	<u>WANG</u>	<u>Baomin</u>	
0	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	DATE: <u>2/27/02</u>
4	POST OFFICE ADDRESS	<u>Rockville</u>	<u>Maryland MD</u>	COUNTRY OF CITIZENSHIP
		<u>GlaxoSmithKline</u>	<u>Research Triangle Park</u>	STATE & ZIP CODE/COUNTRY
		<u>Five Moore Drive, PO Box 13398</u>		<u>NC 27709 US</u>
	FULL NAME OF INVENTOR	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME/INITIAL
	INVENTOR'S SIGNATURE			DATE:
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP
	POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE & ZIP CODE/COUNTRY
	FULL NAME OF INVENTOR	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME/INITIAL
	INVENTOR'S SIGNATURE			DATE:
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP
	POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE & ZIP CODE/COUNTRY
	FULL NAME OF INVENTOR	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME/INITIAL
	INVENTOR'S SIGNATURE			DATE:
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP
	POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE & ZIP CODE/COUNTRY
	FULL NAME OF INVENTOR	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME/INITIAL
	INVENTOR'S SIGNATURE			DATE:
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP
	POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE & ZIP CODE/COUNTRY
	FULL NAME OF INVENTOR	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME/INITIAL
	INVENTOR'S SIGNATURE			DATE:
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP
	POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE & ZIP CODE/COUNTRY
	FULL NAME OF INVENTOR	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME/INITIAL
	INVENTOR'S SIGNATURE			DATE:
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP
	POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE & ZIP CODE/COUNTRY

WO 01/02580

PCT/GB00/02357

## SEQUENCE LISTING

<110> Glaxo Group Limited  
 Anderson, David M  
 Liu, Lin  
 Podkovyrov, Sergey  
 Wang, Baomin

<120> Vectors, cells and processes for pyrimidine  
 deoxyribonucleosides production

<130> P1049

<140>

<141>

<150> US 09/345492

<151> 1999-07-01

<150> US 60/141827

<151> 1999-07-01

<160> 9

<170> PatentIn Ver. 2.1

<210> 1

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 1

tattctagac gatttcaag ttgaggactt atgc

34

<210> 2

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 2

tatatcgata attcattaca atttacacgc tgcac

35

<210> 3

WO 01/02580

PCT/GB00/02357

<211> 35

<212> DNA

<213> Artificial Sequence

5 <220>

<223> Description of Artificial Sequence: Primer

<400> 3

tatatcgata aatgtaaatt taaggattct aaatg

35

10

<210> 4

<211> 34

<212> DNA

<213> Artificial Sequence

15

<220>

<223> Description of Artificial Sequence: Primer

<400> 4

20 tatgtcgact ccttaaaagt atttttaaa actc

34

<210> 5

<211> 30

<212> DNA

25 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: DNA insert

30

<220>

<223> Part of a double stranded DNA insert. The non-overlapping part of SEQ ID NO: 6 is the complement

35

<400> 5

cgagcccgcc taatgagcgg gcttttttt

30

<210> 6

<211> 38

40 <212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: DNA insert

45

<220>

<223> Part of a double stranded DNA insert. SEQ ID NO: 5 is the complement of the non-overlapping part of the sequence

WO 01/02580

PCT/GB00/02357

<220>

<221> misc\_feature

<222> (1)..(4)

5 <223> Overhang

<220>

<221> misc\_feature

<222> (35)..(38)

10 <223> Overhang

<400> 6

gtacaaaaaa aagcccgctc attaggcggg ctcgggcc

38

15 <210> 7

<211> 37

<212> DNA

<213> Artificial Sequence

20 <220>

<223> Description of Artificial Sequence:

Oligonucleotide

<400> 7

25 gatccggagg ataatgaaa caataccaag attaat

37

<210> 8

<211> 31

<212> DNA

30 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:

Oligonucleotide

35

<400> 8

taaattctgg tattgttca ttatcctcc g

31

40 <210> 9

<211> 45

<212> DNA

<213> Artificial Sequence

45 <220>

<223> Description of Artificial Sequence: Primer

<400> 9

agcaaacatt aaacagcgtg caattacata tgataatca gggtc

45